Study Design

Grinnell College

Fall 2025

Review

We have spent time looking at what we can do with data.

- Making graphics + visuals
- Describing graphics + visuals
- Tables

Outline

Statistics (largely) involves the following three broad domains:

- Design – how do we obtain our data
- Description – graphics and summaries
- Inference – decision-making

By the end of today you will be able to answer:

- What is the difference between experiments and observational studies?
- How do we make *causal* claims? (cause and effect)
- How do we avoid *biases* in our data?

Review (again)

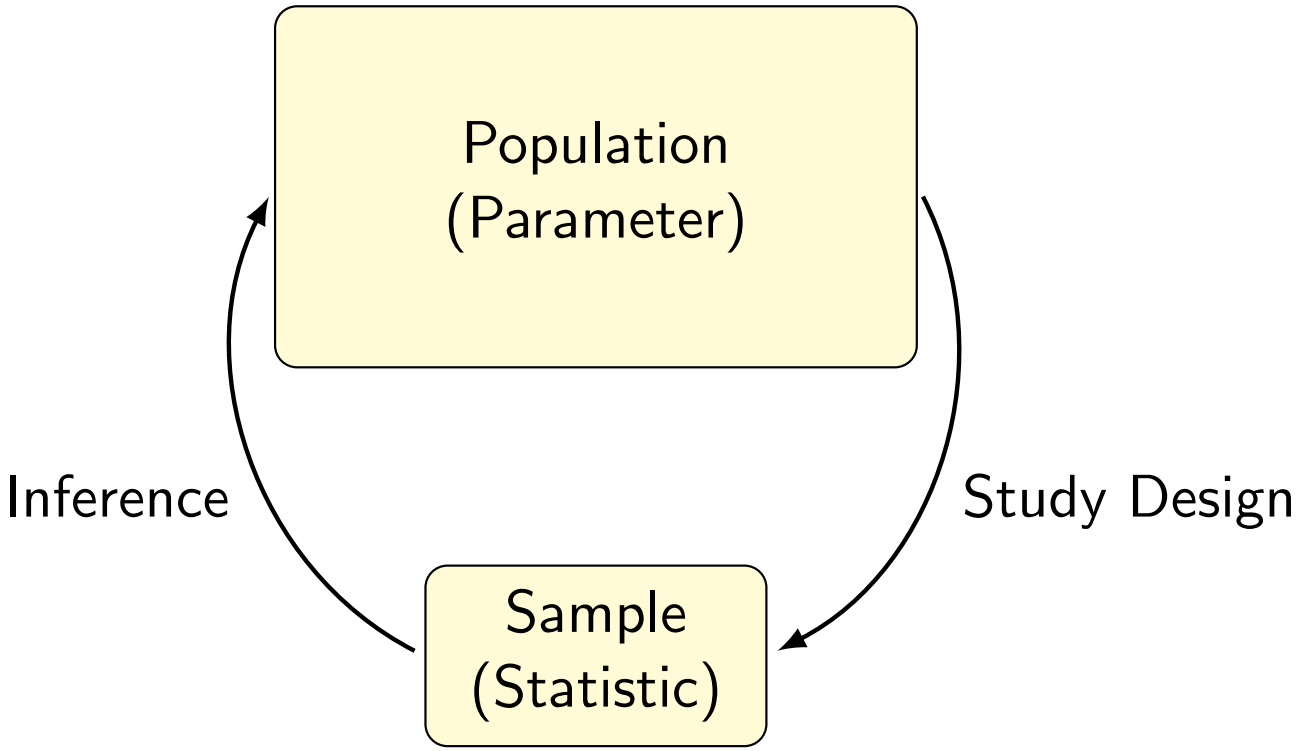
Population is a big group of subjects/events/things about which we wish to learn about

Parameter is a *quantifiable* attribute of a population. Most of the time, the parameter value is unknown

Sample is a much smaller, subgroup of a larger population

Statistic is a numerical summary of the sample that we calculate from our sample data

The Statistical Framework



Anecdotal Evidence

source: IMS Textbook

1) A man on the news got mercury poisoning from eating swordfish, so the average mercury concentration in swordfish must be dangerously high.

2) I met two students who took more than 7 years to graduate from Duke, so it must take longer to graduate at Duke than at many other colleges.

3) My friend's dad had a heart attack and died after they gave him a new heart disease drug, so the drug must not work.

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 need to control everything that could change the response except for the treatments
  - randomization of treatments is the best way to accomplish this

Observational Studies

- No random assignment $\rightarrow$ can't definitively establish cause-and-effect relationships
- we are limited to talking about associations only

Surveys

Surveys are a type of observational study where we ask people about their attitudes/opinions/beliefs

Some important things to think about:

- How do we select people we talk to?
- How many people do we talk to?
- How do we obtain information?
  - phone, email, in-person?

Surveys

How do we select people?

We want our sample to be **representative** of our population.

- this means that our sample is nearly the same as our population, only smaller
  - i.e.: same proportions M/F, same age/ethnic demographics
- a representative sample allows us to generalize our results from the sample to the pop.
- **biased**: a sample that is **not** representative

Not always possible to get a representative sample, but this is what we want to strive for

How do we select people?

Random Sample

We can choose people at random from our population to reduce the chances of getting a biased sample.

- usually the best way to get a representative sample
- allows us to **generalize** from our sample to the pop.

Sample Size (n = ?)

- number of people we survey is important (more people = more info)
- sample size is very important, but proportion of pop. surveyed is not
- better to have small sample that is representative, than a large sample that is biased

How do we select people?

Census – Would conducting a census of the entire population be better?

Issues

- difficult
- time consuming
- expensive

Sampling Methods

Simple Random Sample (SRS)

- each combination of observations has same chance of being selected
- if we were to select multiple SRSs again, different observations will be chosen
  - variability by sampling $\rightarrow$ sampling variability

Stratified Sample

- **strata**: subgroups of pop., within each the individuals are similar
  - individuals within a strata are similar, but strata themselves can be very different
- we take a SRS from each strata to ensure each group has representation

geologylearn.blogspot.com/2015/10/rock-layers.html

Sampling Example

Scenario: We want to ask 200 Grinnell students their opinion on if they prefer living in the dorms

How do we do this with an SRS?

Why might we want to use a stratified sample?

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

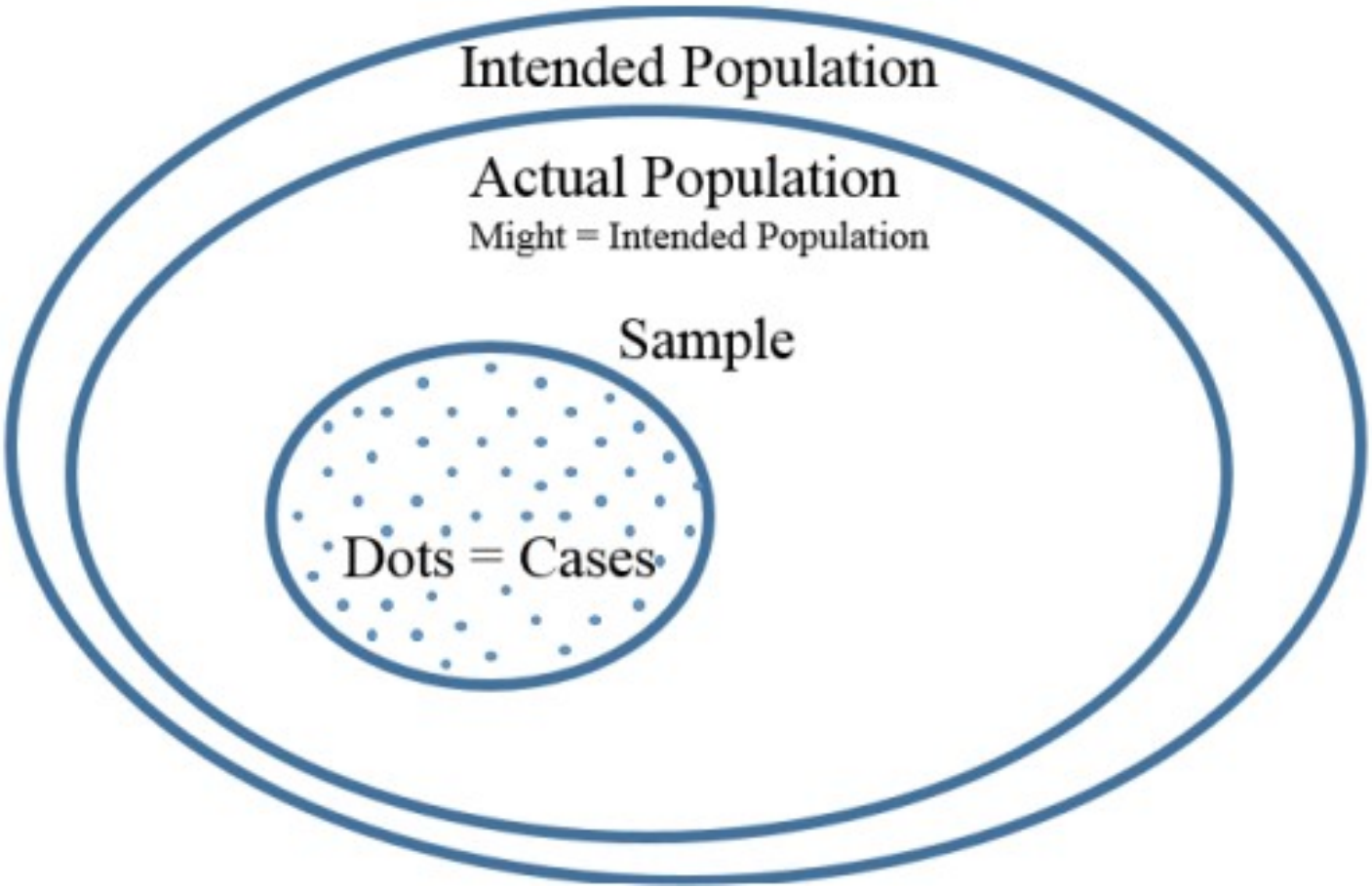
Non-response Bias

- 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

Intended vs. Actual population



source: Dr. Ziegler's Stat 104 notes (ISU)

Observational Studies

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:

- **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

Treatments

- combo of factors and levels that are given 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')

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
- control outside sources that might affect response (other variables)

Replication – 2 types

- having multiple *replicates* (cases = EUs) for each treatment
- *repeatability*: 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)

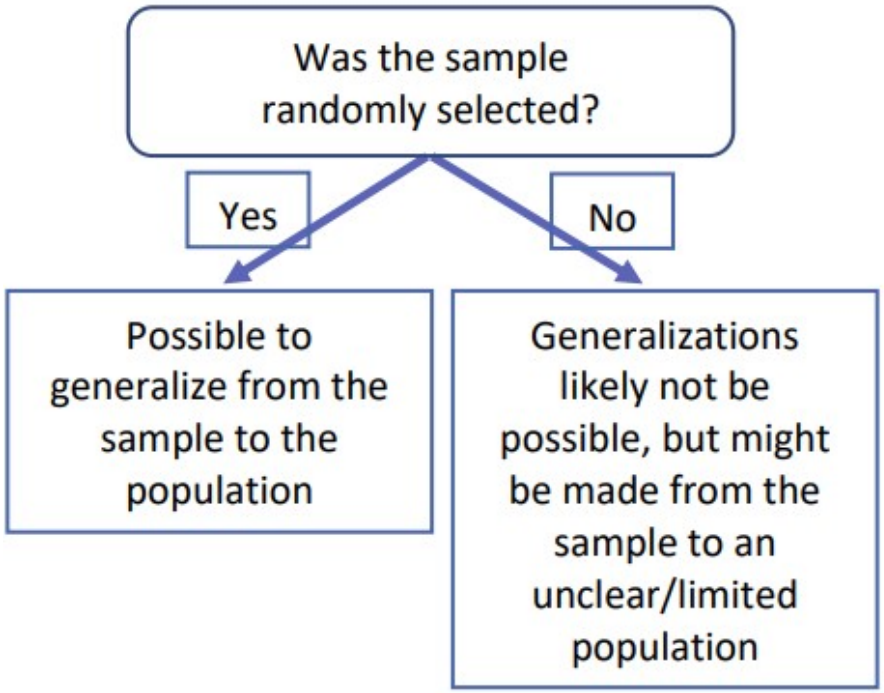
Random Sampling: picking our sample at random from the population

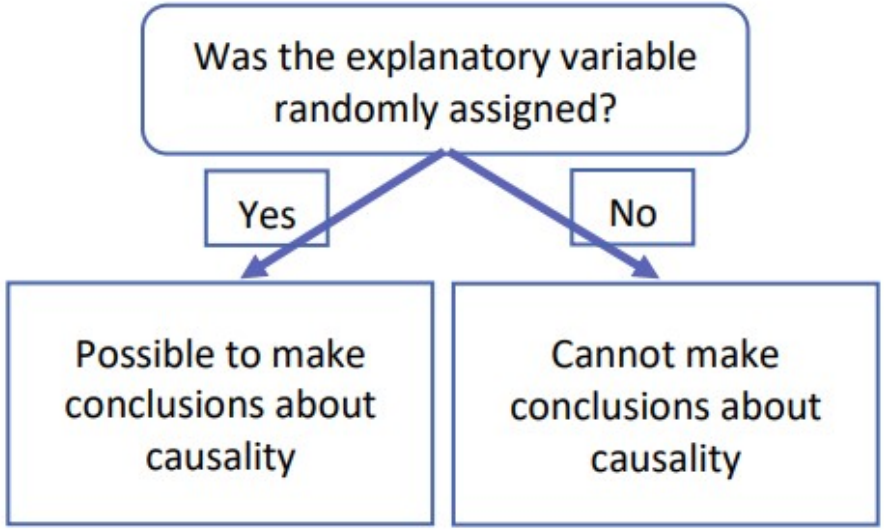
- goal: get sample that is very similar to our population (representative)
- allows us to generalize our conclusions about the sample to the entire population

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

- allows us to make cause-and-effect claims
- balances out effect of confounding variables between both groups so that they don't affect our results
- 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 affects our results when we don't want it to

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

How does Randomization fix this issue?
Random assignment balances a (potentially) confounding variable between the groups $\rightarrow$ it doesn't interfere with out results

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?