

Topic 3: Experiments

STOR 155: Introduction to Data Models and Inference
Dr. Teressa Bergland
Fall 2025





Announcements

Course PSAs:

- Reminder: Tutorial sessions start this week!
- Homework 2 due **TODAY, Tuesday 1/20 on WebAssign**
 - You likely can still complete this homework even if you have not yet paid for WebAssign... but if you haven't yet, it's time to do that.
- Homework 3 due **Thursday 1/22 on WebAssign**
- Attendance: If you are getting errors related to location, try accessing PollEverywhere a different way (different device, app vs browser, etc)



Let's Practice! Choosing Sampling Methods

The native Yup'ik peoples of Alaska have seen an increase in chronic disease over the last several decades, despite little to no change to their traditional lifestyle. The problem is especially bad in rural and isolated areas.

- You are investigating whether these diseases might stem from increased heavy metal concentrations in the traditional Yup'ik diet.
- Your **first step** is to determine whether the Yup'ik have abnormally high concentrations of heavy metals in their blood.
- A local expert informs you that there are approximately 60 rural Yup'ik villages in Alaska, ranging from 50-250 residents each.

How might you collect a sample from this population for your study?



Additional Factors in Causing Bias

Wording

Survey of high-school students:

- “Which is easier for someone of your age to buy: cigarettes, beer, or marijuana?” (35%, 18%, 34%)
- “Which is easier for someone of your age to obtain: cigarettes, beer, or marijuana?” (39%, 27%, 19%)

Poverty Assistance:

- “Is US spending too much on assistance to the poor?” (13%)
- “Is US spending too much on welfare?” (44%)

Framing

Fewer people mention the economy in open-ended version

% answering that the issue matter most in deciding their vote for president in 2008

	Open-ended	Closed-ended
The economy	35	58
The war in Iraq	5	10
Health care	4	8
Terrorism	6	8
Energy policy	*	6
Other	43	8
Candidate mentions	9	–
Moral values/social issues	7	–
Taxes/distribution of income	7	–
Other issues	5	–
Other political mentions	3	–
Change	3	–
Other	9	–
Don't know	7	2
	100	100

Note: Open-ended figures reflect respondents' unprompted first response. Close-ended figures reflect respondents' first choice from five options read by the interviewer.
Source: Survey conducted November 2008.



Sample statistic, Population parameter

Survey of high-school students:

- “Which is easier for someone of your age to buy: cigarettes, beer, or marijuana?” (35%, 18%, 34%)
- “Which is easier for someone of your age to obtain: cigarettes, beer, or marijuana?” (39%, 27%, 19%)

Poverty Assistance:

- “Is US spending too much on assistance to the poor?” (13%)
- “Is US spending too much on welfare?” (44%)

Populations have **parameters**
(unknown true values of interest)

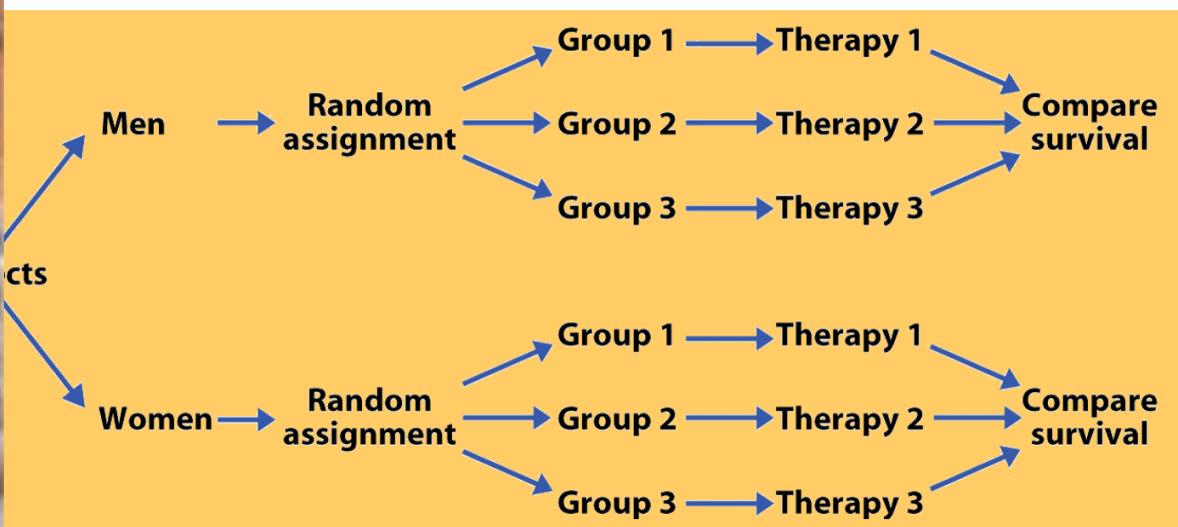
Samples have **statistics**
(observed values from sample)

Goal: a good, unbiased sample
has statistics that are good
estimates of parameters

Experimental Design

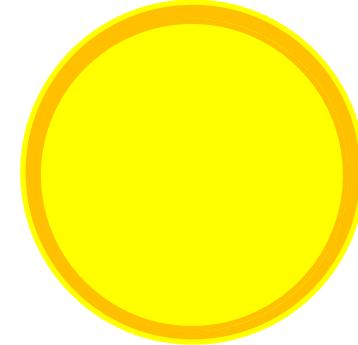


Design:





Let's Practice! Variables



A study is designed to test the effect of **light level** and **air quality** on the **growth rate of corn**. The researcher also believes that light level and air quality might have different effects on corn grown in soil with different acidity levels, so wants to make sure varying **soil acidity levels** are equally represented in each group.

- a) There are 3 explanatory variables (light, air, soil) and 1 response variable (corn growth)
- b) There are 2 explanatory variables (light and air), 1 blocking variable (soil), and 1 response variable (corn growth)
- c) There is 1 explanatory variable (soil) and 3 response variables (light, air, corn growth)
- d) There are 2 blocking variables (light and air), 1 explanatory variable (soil), and 1 response variable (corn growth)



Eliminating Bias

Important terms:

- Placebo
- Placebo effect
- Blinding
- Double-blind



Experiments vs. Observational Studies

Random sampling	Random assignment	No random assignment	Generalizability
ideal experiment	Causal conclusion, generalized to the whole population.	No causal conclusion, correlation statement generalized to the whole population.	most observational studies
most experiments	Causal conclusion, only for the sample.	No causal conclusion, correlation statement only for the sample.	No generalizability
	Causation	Correlation	bad observational studies



Practice Time! Experimental True/False

Which of the following statements regarding differences between experiments and observational studies are true?

- a) Experiments take place in a lab, while observational studies do not need to.
- b) In an observational study we only look at what happened in the past.
- c) Experiments use random assignment while observational studies do not.
- d) Observational studies are completely useless since no causal inference can be made based on their findings.
- e) Experiments involve active intervention or treatment, while observational studies are passive.