Non-Sampling Errors

Week 1

Stat 260, St. Clair

Content:

- Non-sampling errors
- Selection biases
- Measurement errors

Non-sampling Errors

- **Non-sampling Error** Systematic variation between sample data and a population data that can't be explained by sampling error. Can introduce *bias* into results.
    - selection bias
    - measurement error

Selection bias

- **Undercoverage** Occurs when the sampling frame does not cover the targeted population.
  - units in the undercoverage part of the population have **no chance** of being selected for the sample.

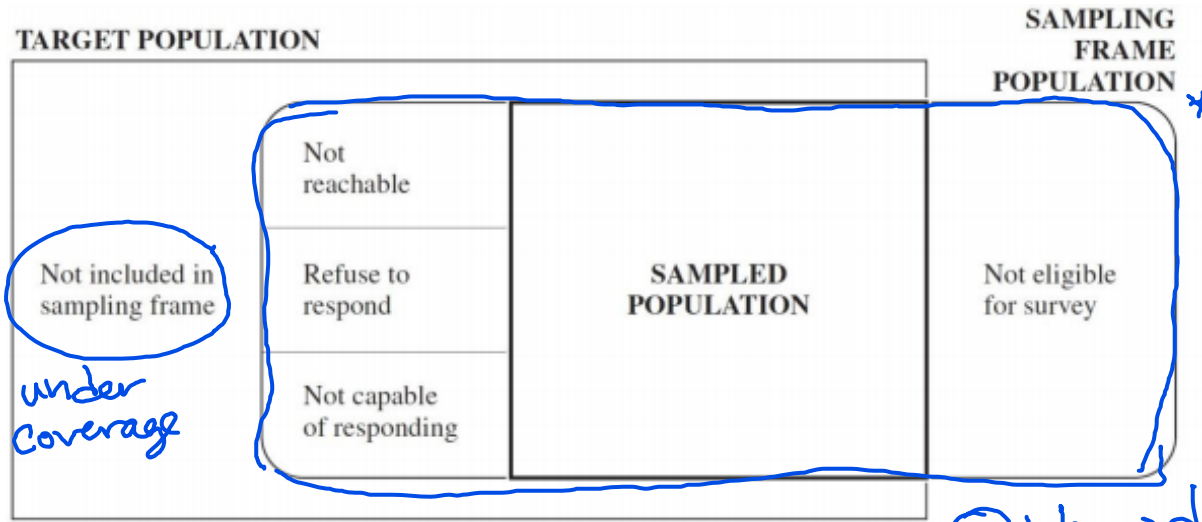

Figure 1: Target population and sampling frame (Lohr, 2010, p. 4)

Selection bias

- **Nonresponse** Failure to collect responses from all observation units selected for a survey.

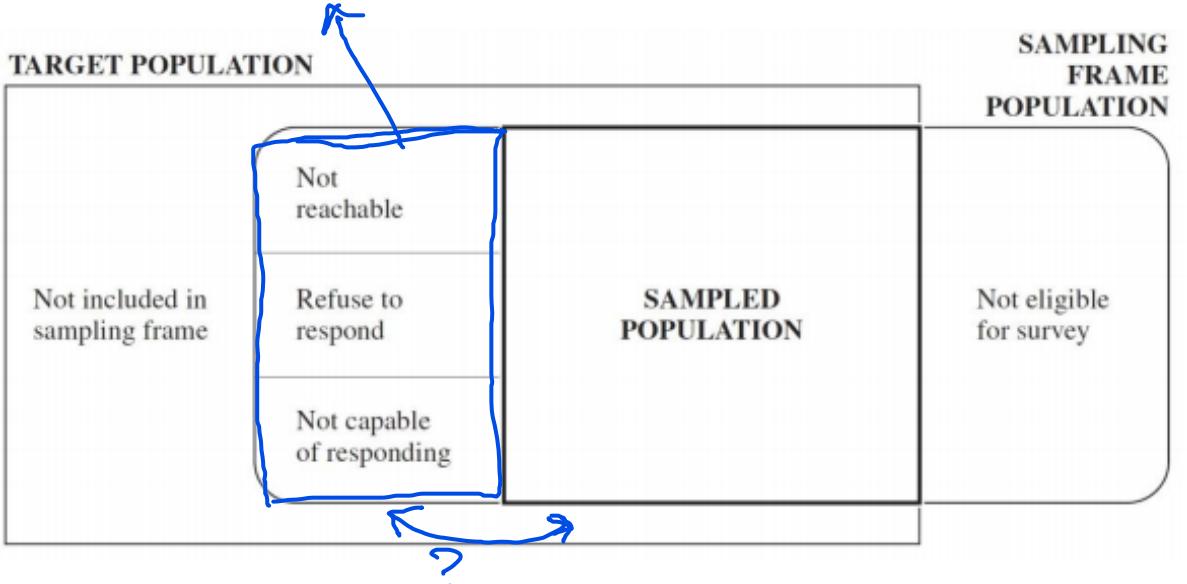

Figure 1: Target population and sampling frame (Lohr, 2010, p. 4)

S.M

S.M. emails → Spam folder

Selection bias

- **Nonresponse**
  - can't contact (no substitutions!!)
  - can't respond (e.g. illness, lack of knowledge)
  - refuse to answer (e.g. privacy, fear, embarrassment)
  - effect not easily measured but nonresponse rates can be reduced by careful planning (callbacks, mailings, incentives, shorten interview length)

"Correcting" for selection biases

- Many large-scale surveys **"weight"** data from respondents to be "more representative" of the target population.
    - More during week 8

Measurement Errors

- **Measurement Instrument Errors** Occur due to
  - inaccurate measurement tools/machines
  - poorly worded questions
  - inappropriate questions/questionnaire design, etc....

Measurement Errors

- In ecology, we often have **detection** errors because we don't detect every animal/plant that we are looking for
    - e.g. moose are hidden by trees and don't get counted
    - a raw count of moose in a plot will **underestimate** the total number of moose in the plot

Measurement Errors

- **Interviewer bias** Occurs because people tend to give different answers to different interviewers.

Measurement Errors

- **Response Errors** People differ in motivation and ability to respond to survey questions.

Questionnaire Design - some topics

Wording of Questions

- Keep questions simple but specific.
- Make sure all terms or measurement amounts are well-defined.

Questionnaire Design - some topics

Wording of Questions

- Don't use leading or loaded questions or words.
- Don't use **double-barreled** questions that deal with 2 or more issues.
    - 2019 Trump tweet: "Wow! A Suffolk/USA Today Poll, just out, states, “50% of Americans AGREE that Robert Mueller’s investigation is a Witch Hunt.”

Questionnaire Design - some topics

Wording of Questions

- Don't use leading or loaded questions or words.
- Don't use **double-barreled** questions that deal with 2 or more issues.
    - 2019 Trump tweet: "Wow! A Suffolk/USA Today Poll, just out, states, “50% of Americans AGREE that Robert Mueller’s investigation is a Witch Hunt.”
    - but here is the question asked: "President Trump has called the Special Counsel’s investigation a ‘witch hunt’ **and** said he’s been subjected to more investigations than previous presidents because of politics. Do you agree?"

https://www.factcheck.org/2019/03/trump-touts-questionable-survey-results/

Questionnaire Design - some topics

Question Ordering

- General vs. Specific: usually best to ask general questions prior to specific questions
- Order of possible responses
    - verbal questions: response more likely to be the last option in the list
    - written questions: response more likely to be the first option in the list
- Solutions: vary ordering, repeat question

Questionnaire Design - some topics

Open vs. Closed Questions

- Open question - respondents not prompted with categories for response. Good for exploratory surveys or sensitive questions but can be hard to analyze.
- Closed question - respondents given a list of possible responses to choose from.

Questionnaire Design - some topics

Response Options for Closed Questions

- Can't list all possible responses but pretesting of questions should narrow the list to the most common responses.
- Have other" category when appropriate.
- Detailed/narrow topic questions: Provide a screening question to assess an individuals knowledge of a specific issue.

Non-Sampling biases vs. estimator biases

- Non-sampling errors can introduce biases into our conclusions.
  - extremely hard to actually quantify
- When a **probability sampling design** is used, we can quantify (or approximate) an **estimator's bias**
  - e.g. *Does this estimation method systematically over- or under-estimate the population parameter?*
  - **Estimator bias** assumes **no non-sampling errors**!!

Ch. 2+