# 36-303: Sampling, Surveys and Society

What is Sampling?
Brian Junker
132E Baker Hall
brian@stat.cmu.edu

---

## Review Quiz

- Fill in your name.
- Answer questions on the handout.
- You have 30 minutes.

(If you are done early read through today's Blog posts handout with care.)

---

## Handouts

- Project Ideas
- Project Schedule
- Topics Schedule
- Quiz
- Lecture Notes
  - Quiz
  - Forming Project Groups
  - Project Ideas and Class Schedule
  - What is Sampling?

---

## Forming Project Groups

- Find people you can work with
- Use the discussion board Blackboard to find a group to join or find a person to add to your group
- Groups should be 4-5 students each
- *Email brian@stat.cmu.edu with your proposed group members, by Friday at 5:00pm. ONE EMAIL PER GROUP.*
- *I will assign you to a group if you do not choose.*

## Project Ideas, Class Schedule

- Project Ideas

- Project Schedule

- Topics Schedule

## What is Sampling?

- Sampling is a statistical process of "purposefully" selecting a subset of units from a population in order to make inferences about the entire population.
  - The "best" way to select a sample is through the use of probability methods, because this gives us a basis for inference.
  - *Utah v. Evans* (2003) : When is it OK to use sampling to supplement the US Census?

## Elements of a Sample

- Key elements to understanding properties of sample include:
  - *Target Population* - collection of observations we want to study (e.g. possible voters in NH).
  - *Sampled Population* – all possible observation units that might have been sampled.
  - *Sampling Frame* – list of all sampling units (e.g. list of telephone numbers.
  - *Sample* - subset of population.
  - *Sampling Unit* - unit we actually sample (e.g. household).
  - *Observational Unit* - element to be measured (e.g. individual).

## Schema

Target Population
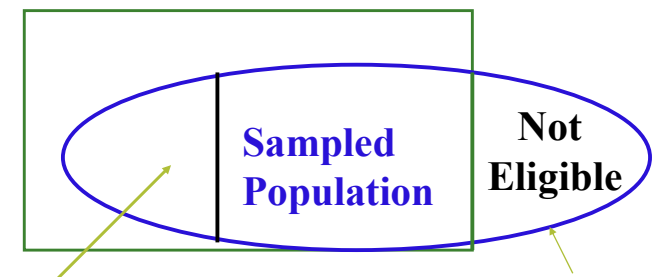


**Refusals,Unreachable, Incapable, etc.**

Sampling Frame Population

# Does Sample Represent Population?

- "Representativeness" comes from
  (a) match between target population and sampled population.
  (b) method for drawing sample.
- Two kinds of errors:
  - Non-sampling – *can be reduced by careful design of the survey*
  - Sampling – *can be quantified by statisitcs, reduced by increasing sample size*

# Two Kinds of Errors

- Non-sampling errors:
  - Selection bias - part of target population is not in sampled population.
  - Measurement bias - measuring instrument has tendency to differ from true value in one direction.
- Sampling error **-** results from taking a sample instead of whole population.

# Why the polling discrepancies?

- Sampling errors?
  - What was the sample size?
  - How do you calculate the "error"?
- Non-sampling errors?
  - Selection bias

  - Measurement bias

# Methodological Features of Examples

- What can we say about:
  - population of interest
  - frame/list
  - sampling technique
  - sample size
  - response rate
  - mode of interview
  - possible sources of selection bias and inaccuracy
  - other details on methodology relevant to inferences of interest

# Summary of Today's Lecture

- Review Quiz.
- Project groups
- Project proposals
- Key elements of sampling
- What makes a sample representative?