

**Introduction to Survey Sampling**

**Karen Foote Retzer**

Survey Research Laboratory  
www.srl.uic.edu

1

---

---

---

---

---

---

---

---


**Census or sample?**

**Census:**

- Gathering information about every individual in a population

**Sample:**

- Selection of a small subset of a population

Survey Research Laboratory 

2

---

---

---

---

---


---

---

---

**Why sample instead of taking a census?**

- Less expensive
- Less time-consuming
- More accurate
- Samples can lead to statistical inference about the entire population

Survey Research Laboratory 

3

---

---

---

---

---

---

---

---

## Probability vs. non-probability sample

### Probability sample:

- Generalize to the entire population
- Unbiased results
- Known, non-zero probability of selection

### Non-probability sample:

- Exploratory research
- Convenience
- Probability of selection is unknown

Survey Research Laboratory 

4

---

---

---

---

---

---


---

---

## Target population

**Definition:** The population to which we want to generalize our findings.

- **Unit of analysis:** Individual/household/city
- **Geography:** State of Illinois/Cook County/Chicago
- **Age/gender**
- **Other variables**

Survey Research Laboratory 

5

---

---

---

---

---

---

---

---

## Examples of target populations

- Population of adults (18+) in Cook County
- UIC faculty, staff, students
- Youth age 5–18 in Cook County

Survey Research Laboratory 

6

---

---

---

---

---

---

---


---

## Sampling frame

A complete list of all units, at the first stage of sampling, from which a sample is drawn.

**Examples:**

- Lists of people or schools
- Phone numbers in specific area codes
- Maps of geographic areas

Survey Research Laboratory 

---

---

---

---

---

---

---

---

## Sampling frames (cont'd)

**Example 1:**


- **Population:** Adults (18+) in Cook County
- **Possible frames:** list of phone numbers, list of block maps, list of addresses

**Example 2:**

- **Population:** Females age 40-60 in Chicago
- **Possible frame:** list of phone numbers, list of block maps

**Example 3:**

- **Population:** Youth age 5-18 in Cook County
- **Possible frame:** List of schools

Survey Research Laboratory 

---

---

---

---

---


---

---

---

## Sample designs for probability samples

- Simple random samples
- Systematic samples
- Stratified samples
- Cluster

Survey Research Laboratory 

---

---

---

---

---


---

---

---

## Simple random sampling

- **Definition:** Every element has the same probability of selection *and* every combination of elements has the same probability of selection.
- **Probability of selection:**  $n/N$ , where  $n$ =sample size;  $N$ = population size
- Use random number tables, software packages to generate random numbers
- Most precision estimates assume SRS.

Survey Research Laboratory 

10

---

---

---

---

---


---

---

---

## Systematic sampling

- **Definition:** Every element has the same probability of selection, but not every combination can be selected.
- Use when drawing SRS is difficult.
  - List of elements is long & not computerized
- **Procedure:**
  - Determine population size  $N$  & sample size  $n$
  - Calculate **sampling interval** ( $N/n$ )
  - Pick random start between 1 & sampling interval
  - Take every  $i$ th case
  - Problem of **periodicity**

Survey Research Laboratory 

11

---

---

---

---

---

---


---

---

## Stratified sampling

*Proportionate*

- **To ensure sample resembles some aspect of population**
- **Population is divided into subgroups (strata)**
  - Students by year in school
  - Faculty by gender
- **Simple random sample (with same probability of selection) taken from each stratum.**

Survey Research Laboratory 

12

---

---

---

---

---

---


---

---

## Stratified sampling (cont'd)

*Disproportionate*

- Major use is comparison of subgroups
- Population is divided into subgroups (strata)
  - Compare boys & girls who play Little League
  - Compare seniors & freshmen who live in dorms
- Probability of selection needs to be higher for smaller stratum (girls & seniors) to be able to compare subgroups.
- Post-stratification weights

Survey Research Laboratory 

13

---

---

---

---

---


---

---

---

## Cluster sampling

- Typically used in face-to-face surveys
- Population divided into clusters
  - Schools (earlier example)
  - Blocks
- Reasons for cluster sampling
  - Reduction in cost
  - No satisfactory sampling frame available

Survey Research Laboratory 

14

---

---

---

---

---


---

---

---

## Determining sample size: SRS

- Need to consider
  - Precision
  - Variation in subject of interest
- Formula
  - Sample size  $n_o = \frac{CF^2 * (pq)}{\text{Precision}}$
  - For example:  $n_o = \frac{1.96^2 * (.5*.5)}{.05^2}$
- Sample size *not* dependent on population size.

Survey Research Laboratory 

15

---

---

---

---

---

---

---

---

## Sample size: Other issues

- Finite Population Correction  $n = n_0 / (1 + n_0 / N)$
- Design effects
- Analysis of subgroups
- Increase size to accommodate nonresponse
- Cost

---

---

---

---

---

---

---

---