

Sampling

February 11, 2004

What is sampling?

- A sample is a small number of subjects chosen to represent the much larger population of interest.
- Sampling is the process for choosing those subjects.

Goal is representativeness

- Sample should be much like the broader population
- The sample is studied or surveyed and results are generalized more broadly

How is sampling done?

- Random samples
 - *Simple*
 - *Stratified*
 - *Cluster*
- Non-random
 - *Systematic*
 - *Convenience*
 - *Purposive*

These are the two broad approaches to sampling

Simple random sample

- Each and every member of the population has an equal chance of selection
- Selection is not based on another's selection - it is independent

Simple random technique

- Table of random numbers
 - *Available in statistics textbooks*
- Computer software or on the web
- Advantages
 - *If large enough, representativeness is quite possible*
- Disadvantages
 - *Difficult to do; identifying every member of the population is hard*
 - *Can hand the researcher a non-representative sample, so it's not used if researchers want to ensure representativeness*

An example

- EDF 6481 presentations near end of term
- All students are assigned a number
- We use a table or random number generator to determine the order of presentations
- We'll use the Research Randomizer linked from the class website

Here are your numbers

Barnes	1	Lands	17
Begley	2	Lowery-Whitney	18
Biringer	3	Macwithney	19
Burton	4	Marin	20
Carter	5	McCallister	21
Conrad	6	Mulder	22
Danesh	7	Mustakas	23
Dettman	8	Ni	24
Dube	9	Saenz	25
Feury	10	Schellhase	26
Fortuno	11	Schneider	27
Goodin	12	Snipes	28
Gott	13	Spazante	29
Hollenberger	14	Spiller	30
Johnson	15	Tafur	31
King	16	Yount	32

Stratified random technique

- Identify subgroups (i.e. males/females, or teachers/students)
- Draw random samples from each subgroup
- Advantages
 - More likely to result in representativeness, especially with small samples
- Disadvantages
 - Researcher has a more complicated task

An example

- I want to survey all 400 8th graders
- I can administer the survey to 200 students
- I believe SES is relevant
- Free/Reduced lunch data
 - No lunch subsidy 50% 100
 - Reduced price lunch 35% 70
 - Free lunch 15% 30
- For my sample, I draw 100, 70, and 30 from respective groups

Cluster random sampling

- Used when individuals can't be chosen
 - This is often the case in schools
- Much like simple random sampling, except clusters are randomly chosen
 - I.e. certain schools in the state
- Advantages
 - Easy
- Disadvantages
 - Not always representative
- Rookie Alert! - This technique takes more than one cluster

Systematic sampling

- This is not random
- Uses a random start and chooses every nth subject to be a member of the sample
- Example: Every 10th student out of 500
- Be certain the list is not ordered in a way that will effect this sort of sampling - perhaps randomize the list

Convenience sample

- Sometimes the population is hard to reach, and researchers must take what they can get
- This is a convenience sample
- Easy to do, but may not be representative
- When used, provide lots of descriptive information

Purposive sampling

- Personal judgment is used to select the sample
- Used when one believes these subjects have the desired knowledge about the population

How large a sample?

- There are fancy ways to calculate this
- A general guideline is to use the size of sample you can within reasonable limits of time, money, and other resources