Statistical Sampling

A guide for gathering data

Tabel of Contents

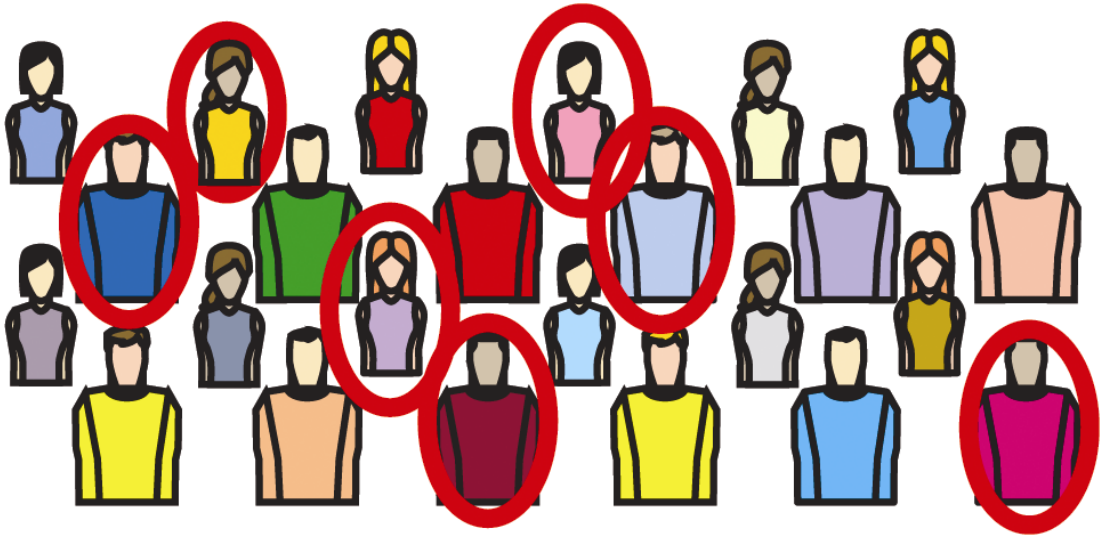
Objectives

- Define the five basic sampling methods
  - Random
  - Systematic
  - Stratified
  - Cluster
  - Convenience
- Identify sampling methods in an example
- Use sampling methods to choose data

TOC

Random Sampling

- The “pick a name out of the hat” technique
  - Random number table
  - Random number generator



Hawkes and Marsh (2004)

TOC

Systematic Sampling

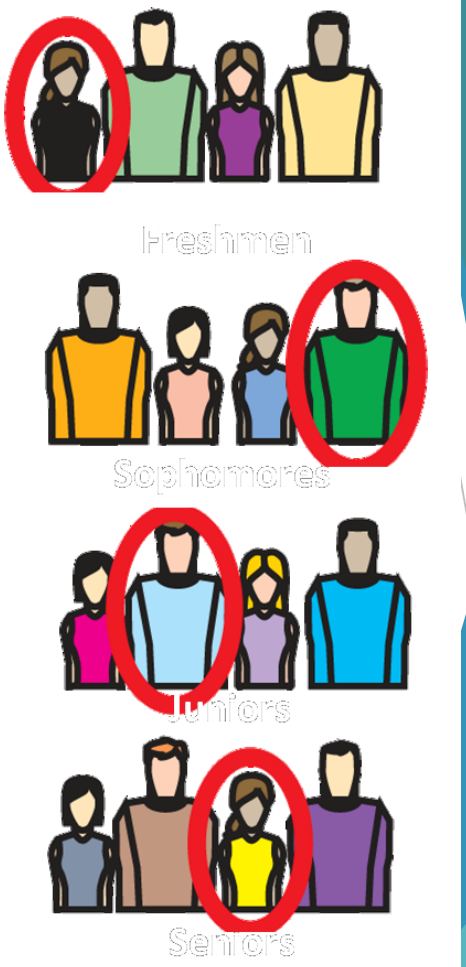
- All data is sequentially numbered
- Every *n*th piece of data is chosen

Hawkes and Marsh (2004)

TOC

Stratified Sampling

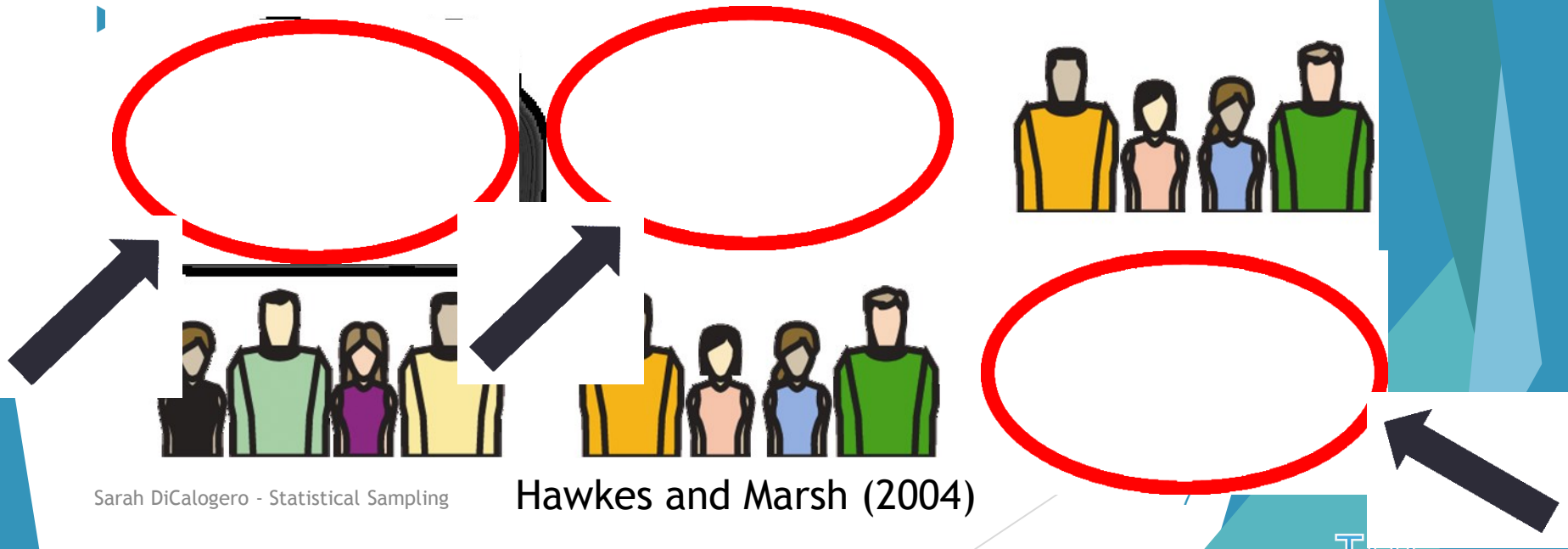
- Data is divided into subgroups (strata)
- Strata are based specific characteristic
  - Age
  - Education level
  - Etc.
- Use random sampling within each strata



Hawkes and Marsh (2004)

Cluster Sampling

- Data is divided into clusters
  - Usually geographic
- Random sampling used to choose clusters



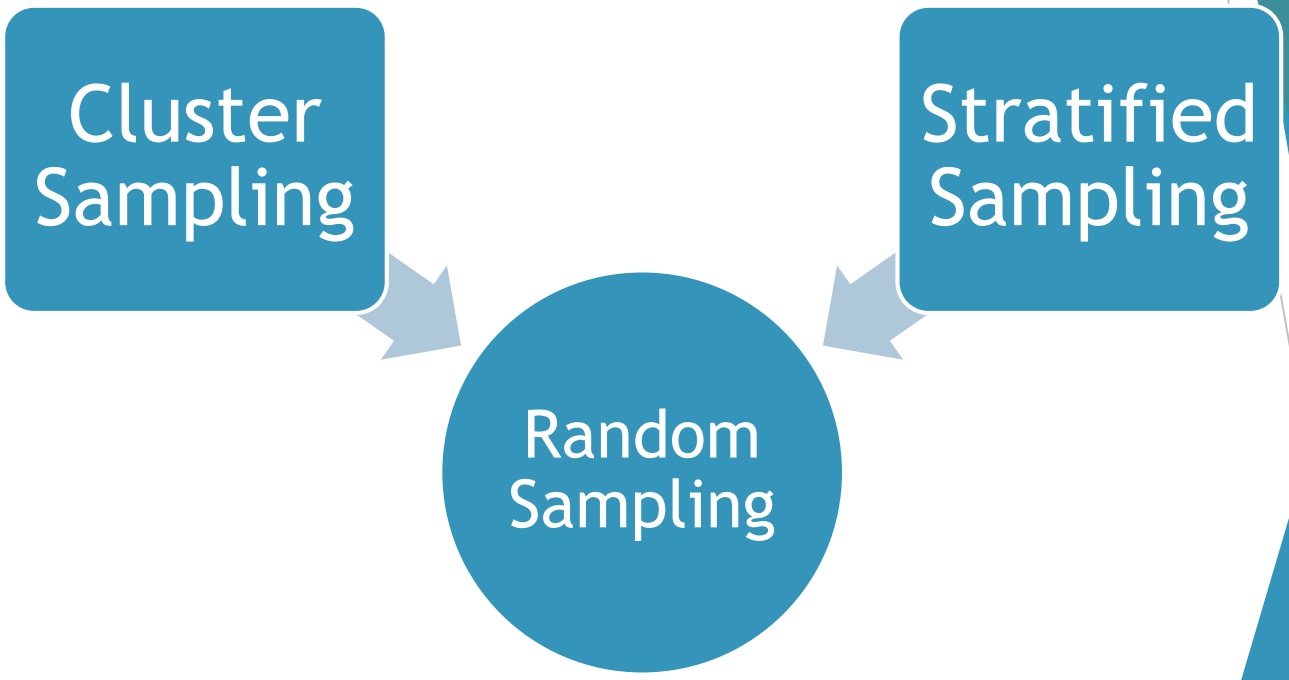
Hawkes and Marsh (2004)

Convience Sampling

- Data is chosen based on convenience
  - BE WARY OF BIAS!

Hawkes and Marsh (2004)

Sampling Relationships



Example 1: Sampling Methods

In a class of 18 students, 6 are chosen for an assignment

Sampling Type	Example
Random	Pull 6 names out of a hat
Systematic	Selecting every 3rd student
Stratified	Divide the class into 2 equal age groups. Randomly choose 3 from each group
Cluster	Divide the class into 6 groups of 3 students each. Randomly choose 2 groups
Convenience	Take the 6 students closest to the teacher

Example 2: Utilizing Sampling Methods

- Determine average student age
  - Sample of 10 students
  - Ages of 50 statistics students

18	21	42	32	17	18	18	18	19	22
25	24	23	25	18	18	19	19	20	21
19	29	22	17	21	20	20	24	36	18
17	19	19	23	25	21	19	21	24	27
21	22	19	18	25	23	24	17	19	20

Example 2 – Random Sampling

- Random number generator

(www.random.org)

Data Point Location	Corresponding Data Value
35	25
48	17
37	19
14	25
47	24
4	32
33	19
35	25
34	23
3	42
Mean	25.1

TOC

Example 2 – Systematic Sampling

- Take every 5^{th} data point

Data Point Location	Corresponding Data Value
5	17
10	22
15	18
20	21
25	21
30	18
35	21
40	27
45	23
50	20
Mean	20.8

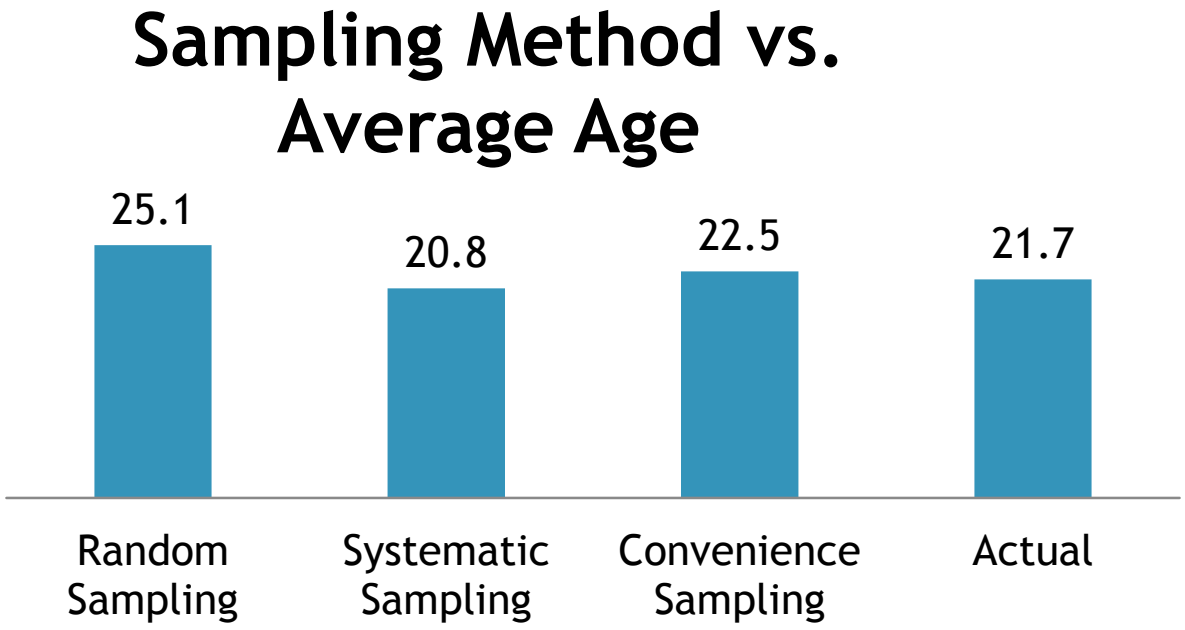
Example 2 – Convenience Sampling

- Take the first 10 data points

Data Point Location	Corresponding Data Value
1	18
2	21
3	42
4	32
5	17
6	18
7	18
8	18
9	19
10	22
Mean	22.5

TOC

Example 2 - Comparison



References

- Hawkes, J., & Marsh, W. (2004). *Discovering Statistics* (2nd ed.). Charleston, SC: Hawkes Publishing Inc..
- YouTube. (2009). *Random Sample*. Retrieved from http://www.youtube.com/watch?v=xh4zxC1OpiA&feature=related
- YouTube. (2009). *Types of Random*. Retrieved from http://www.youtube.com/watch?v=wUwH7Slfg9E&feature=related