Sampling

Dr. Supaporn Erjongmanee

Department of Computer Engineering
Kasetsart University
fengspe@ku.ac.th

Sampling

- Process of collecting subset of population
- Can be divided into
  - Probability sampling
  - Non-probability sampling
- Factors to consider for sampling process
  - Time
  - Place
  - People
  - Cost
  - Sample size
  - And more

Probability Sampling

- Random sampling
  - Each member has equal probability to be chosen
    - Population size = N
    - Each member of population can be selected with probability = 1/N
  - Pro: Give unbiased representation of population
  - Con:
    - When population is large, it is difficult to collect the list "all" members in population

3

Probability Sampling (cont.)

- Systematic sampling
  - Select every kth record from the population
    - k = population size (N) / sample size (n)
  - Similar to random sampling but simpler in practice

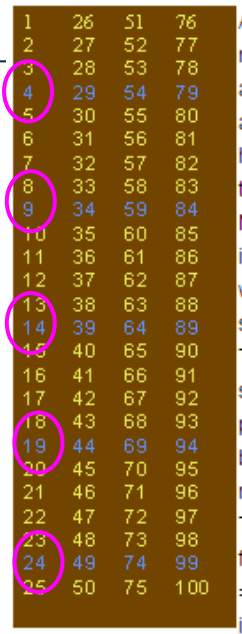
N = 100

want n = 20

N/n = 5

select a random number from 1-5: chose 4

start with #4 and take every 5th unit

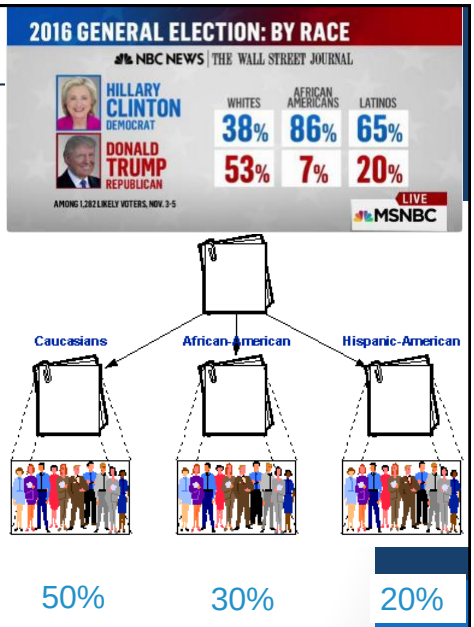


Source: http://www.socialresearchmethods.net/kb/sampprob.php

4

Probability Sampling (cont.)

- Stratified sampling
  - Select according to many sub-groups (strata) first
    - After that, perform random sampling inside each sub-group
  - Sample set holds the same proportion as sub-group proportion
  - If sub-groups = area, also call cluster sampling
  - Pro: Can obtain good representative of samples
  - Con: time consuming



Source: http://www.socialresearchmethods.net/kb/sampprob.php

Probability Sampling (cont.)

- Multi-stage sampling
  - Perform sampling by combining many sampling methods
  - For example: sampling students
    1. Sample schools by districts
       - Stratified sampling
    2. Within districts, sample schools by educational level (elementary vs. high school)
       - Stratified sampling
    3. Within school, random sample students

Source: http://www.socialresearchmethods.net/kb/sampprob.php

Non-Probability Sampling

- Convenience sampling
  - Select any member that is convenient
  - Pro: easy and quick
  - Con: can get biased group of representations

Non-Probability Sampling (cont.)

- Purposive sampling: Sampling with target group
  - Quota sampling:
    - Set quota of samples: e.g. 40 female and 60 male
    - Set minimum of samples: e.g. at least 40 female
  - Snowball sampling
    - Used when desired samples are rare
    - Approach initial subject to generate next subjects
  - Expert sampling
    - Use experts as subjects
  - Modal instance sampling
    - Select typical subjects: average scores, average salary

Example: Finding group of patients with rare disease

Sources: http://www.statpac.com/surveys/sampling.html, http://www.socialresearchmethods.net/kb/sampnon.php

References

1. http://www.socialresearchmethods.net/kb/sampnon.php
2. http://www.statpac.com/surveys/sampling.htm
3. http://www.simplypsychology.org/sampling.html