

Chapter 7. Sampling

- Two types of sampling methods
- Nonprobability sampling
 - Reliance on available subjects
 - Judgmental sampling
 - Snow-ball sampling
 - Quota sampling
- Probability sampling principles
- Probability sampling methods
 - Simple random sampling
 - Systematic sampling
 - Stratified sampling
 - Multistage cluster sampling

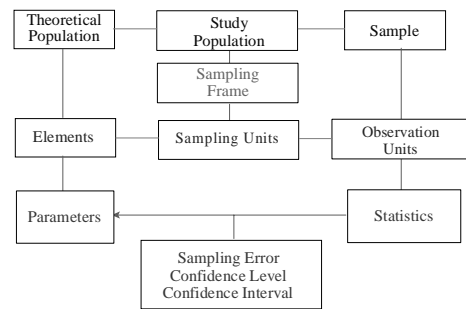
What are the two types of sampling methods?

- Probability sampling: selection of “random” sample. In the sense that every observation in the population has an equal chance to be selected.
 - This is the desirable sampling method because it provides precise statistical descriptions of large populations.
- Nonprobability sampling: when probability sampling are not feasible. This is the less desirable method, but nevertheless commonly used because of practical difficulties with using probability sampling.
 - Nonprobability sampling cannot guarantee that the sample observed is representative of the whole population.

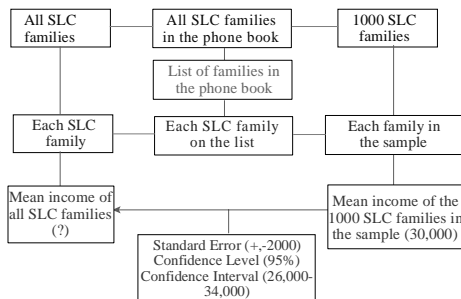
What are the types of nonprobability sampling?

- Reliance on available subjects
 - Examples: Stop people at the mall, University student sample
 - Problems: no sample representativeness
- Purposive or judgmental sampling
 - Examples: friends, colleagues, community leaders
 - Usually used for preliminary testing of questionnaire, and field research
- Snowball sampling
 - Ask people to introduce researcher to more people for interviews
- Quota sampling
 - Step 1. Creating quota matrix: Ex. Gender and age
 - Step 2. Decide on # of observations needed in each quota
 - Step 3. Find subjects with these characteristics to form the sample.

What are the concepts and terminology in probability sampling?



An Example



Concepts continued

- Theoretical population
 - The theoretically specified aggregation of study elements.
- Elements
 - The unit about which information needs to be gathered. Elements provide the basis of analysis.

Concepts continued

- Study population - the aggregation of elements from which the sample is actually selected.
- Sampling frame - the actual list of sampling units from which the sample is selected.

Concepts continued

- Sample
 - A selected group of elements of the study population about which information is gathered.
- Sampling unit
 - The element considered for selection of sampling
- Observation unit
 - An element from which information is collected

Concepts continued

- Parameter
 - A summary description of a given variable in a population
- Statistic
 - A summary description of a given variable in a sample
- Sampling error
 - When using statistics to estimate parameters, the estimation is seldom exact. The error margin is called sampling error. (If you are familiar with the concept of standard error)

Probability Sampling Theory and Sampling Distribution

- The ultimate purpose of sampling is
 - To select a set of elements from a population in such a way that descriptions of those elements accurately portray the parameters of the total population from which the sample is selected.
- Probability sampling
 - Can improve our chance to achieve this goal.
 - Is a very well developed sampling method backed up by probability theory.

Probability Sampling Theory

- Random selection
 - Each element has an equal chance of being selected.
- Reasons of random selection
 - Avoid conscious or unconscious bias by the researcher.
 - Offer access to probability theory, which provides the basis for estimation of parameter and estimation error.

Basic Rules of Sampling

- The larger the sample size, the better chance we have to get an accurate estimate of the population parameter.
- The more homogeneous the population, the smaller the sampling error.

Types of Probability Sampling Designs

- Example research question:
 - We want to select a representative sample of Sichuan Normal University students.
- The goal of sampling is to randomly select 100 students from a total of 20,000 students.
 - Population size = 20,000
 - Sample size = 100

Simple Random Sampling

- Procedure:
 - Step 1. Assign a number between 1 to 20,000 to each student.
 - Step 2. Use the random number table to pick 100 five-digit numbers (random number table on next slide).
 - Step 3. Students whose numbers are picked are in the sample.

10480 15011 01536 02011 81647 91646 69179 14194 62590 36207 20969 99570 91291 90700
 22368 46573 25595 85393 30995 89198 27982 53402 93965 34095 52666 19174 39615 99505
 24130 48360 22527 97265 76393 64809 15179 24830 49340 32081 30680 19655 63348 58629
 42167 93093 06243 61680 07856 18376 39440 53537 71341 57004 00849 74917 97758 16379
 37570 39975 81837 16656 06121 91782 60468 81305 49684 60672 14110 06927 01263 54613

77921 06907 11008 42751 27756 53498 18602 70659 90655 15053 21916 81825 44394 42880
 99562 72905 56420 69994 98872 31016 71194 18738 44013 48840 63213 21069 10634 12952
 96301 91977 05403 07972 18876 20922 94595 56869 69014 60045 18425 84903 42508 32307
 89579 14342 63661 10281 17453 18103 57740 84378 25331 12566 58678 44947 05585 56941
 85475 36857 53342 53988 53060 59533 38867 62300 08158 17983 16439 11458 18593 64952

28918 69578 88231 33276 70997 79936 56865 05859 90106 31595 01547 85590 91610 78188
 63553 40961 48235 03427 49626 69445 18663 72695 52180 20847 12234 90511 33703 90322
 09429 93969 52636 92737 88974 33488 38320 17617 30015 08272 84115 27156 30613 74952
 10365 61129 87529 85689 48237 52267 67689 93394 01511 26358 85104 20285 29975 89868
 07119 97336 71048 08178 77233 13916 47564 81056 97735 85977 29372 74461 28551 90707

51085 12765 51821 51259 77452 16308 60756 92144 49442 53900 70960 63990 75601 40719
 02368 21382 52404 60268 89368 19885 55322 44819 01188 65255 64835 44919 05944 55157
 01011 54092 33362 94904 31273 04146 18594 29852 71585 85030 51132 01915 92747 6-4951
 52162 53916 46389 58586 23216 14513 83149 98736 23495 64350 94738 17752 35156 35749
 07056 97628 33787 09998 42698 06691 76988 13602 51851 46104 88916 19509 25625 58104

48663 91245 85828 14346 09172 30168 90229 04734 59193 22178 30421 61666 99904 32812
 54164 58492 22421 74103 47070 25306 76468 26384 58151 06646 21524 15227 96909 44592
 32639 32363 05597 24200 13363 38095 94342 28728 35806 06912 17012 64161 18296 22851
 29334 27001 87637 87308 58731 00256 45834 15398 46557 41135 10367 07684 36188 18510
 02488 33062 28834 07351 19731 92420 60952 61280 50001 67658 32586 86679 50720 94953

Systematic Sampling

- Easier and as good as simple random sampling if the population is truly randomly ordered.
- Procedure:
 - Step 1. Assign a number between 1-20,000 to each student
 - Step 2. Compute sampling interval
 - Sampling interval = population size / sample size
 - In this example, sampling interval = 20,000/100=200
 - Step 3. Randomly select a beginning number between 1-200, say 43
 - Step 4. Students whose numbers are 43, 243, 443, ... are in the sample

Stratified Sampling

- Can obtain a greater degree of representativeness and decrease the sampling error when used appropriately. This is because creating homogeneous subpopulations can reduce sampling error.
- Earlier conclusion: The more homogenous the population the smaller the sampling error.
- Procedure:
 - Step 1. Choose stratification variables (should be known to you before you do the sampling).
 - Example: class, if there is a reason to believe that students in the same year of college are more alike
 - Step 2. Create Strata.
 - Example: freshmen(6000), sophomore(5000), junior(5000), senior(4000)
 - Step 3. Select a random sample in each stratum. Use either simple random sampling or systematic sampling.
 - Example: in the final sample, freshmen(30), sophomore(25), junior(25), senior(20)

Multistage Cluster Sampling

- Use PPS (Probability proportionate to size sampling) sampling unless the size of each cluster is about the same.
- Do not have to have a complete list of the population.
- Procedure:
 - Step 1. Start with colleges. Estimate the size of population in each college.
 - 10 colleges.
 - Population size in each college: 1000, 1000, 1000, 1000, 2000, 2000, 2000, 3000, 3000, 4000.

- Step 2. Decide the chance of each college to be selected.
- College 1 $1000/20000=5\%$ =>Assign numbers 1-5
- College 2 $1000/20000=5\%$ =>Assign numbers 6-10
- College 3 $1000/20000=5\%$ =>Assign numbers 11-15
- College 4 $1000/20000=5\%$ =>Assign numbers 16-20
- College 5 $2000/20000=10\%$ =>Assign numbers 21-30
- College 6 $2000/20000=10\%$ =>Assign numbers 31-40
- College 7 $2000/20000=10\%$ =>Assign numbers 41-50
- College 8 $3000/20000=15\%$ =>Assign numbers 51-65
- College 9 $3000/20000=15\%$ =>Assign numbers 66-80
- College 10 $4000/20000=20\%$ =>Assign numbers 81-100
- Step 3. Select several colleges.
- Each college's chance to be selected should be proportionate to its size (PPS). Random, systematic, or stratified sampling can be used to select several colleges.
- For example: select five colleges 2,4,7,9,10.
- Step 4. Obtain a list of student names from these selected colleges.
- Step 5. Select 20 students from each selected college.
- Use either simple random, systematic or stratified sampling.

Additional Things to Do

- Take a look at all 7 research articles and see what kind of sampling methods were used in each of them. You will note many of them use convenience samples. Think about why the researchers used these sampling methods, the pros and cons of each, and how you would do things differently if you had all the resources.
- Article links: [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#)