SAMPLE AND SAMPLING TERMINOLOGY

Sample

- A subset, or some part, of a larger whole.
- Larger whole could be anything – bucket of water, a bag of sugar, a group of organizations, a group of students, a group of customers, a group of mid-level managers.

Why sample?

1. Saves cost, labor, and time

- To go for sample study is pragmatic.
- In case population is extremely small, then go for total study. Census another word – total enumeration

2. Quality Management

- Professional fieldworkers – a scarce commodity.
- Instead of doing on large population with less qualified staff, do a sample study with quality fieldworkers.
- Easier to manage small group – quality control. Training, supervision, record keeping.

3. Accurate and Reliable Results

- Properly selected samples are accurate.
- Homogeneous population – only a small sample needed. Likely to be representative. Blood samples.
- Large pop. More non-sampling errors – interviewer mistakes, tabulation errors. Lo quality supervision.

4. No Alternative but Sampling

- For quality control testing may require the destruction of the items being tested e. g. Firecrackers, testing the life a bulb, Testing missiles.
- This is destructive testing.

5. Determine the Period of Study

- Census study requires long time, may be a year. Seasonal variation. For example, Study of unemployment rate over a year. Results refer to which part of the year.

6. Determine the Confidence Level

- Calculate the sampling error – help in determining the confidence level in the data.
- Sampling type may facilitate the use of powerful statistical tests for analysis.

Sampling Terminology

- Number of technical terms used that need explanation.

1. Element

- Unit about which information is collected and is the basis of analysis. Can be a person, a group, a family, an organization, a community.

2. *Population*

- Theoretically specified aggregation of study elements.
- Translating the abstract concept into workable concept. College students. Theoretical explanation.
- Pool of all available elements is population.

3. Target Population

- Out of conceptual variations, what exactly is the focus.
- Complete group of specific population elements relevant to project.
- Call it *Survey population* – aggregation of elements for selecting a sample.
- e.g. study of college students – college students from Govt. institutions, studying social sciences, aged 19 years, and with rural background

4. Sampling

- Process of using a small number of items. Estimate unknown characteristics of population.
- Process of selection – Depending upon the type of sample to be used.

5. Sampling Frame

- List of population elements. Listing of all college students meeting the criteria.
- Also called as *working population* – list that can be worked with operationally. Prepare the list of relevant college students.

6. Sampling Unit

- That element or set of elements considered for selection in some stage of sampling.
- Sampling can be single stage or multistage. Simple or complex.
- In single stage, sampling units are the same as elements.
- In multistage, different levels of sampling units may be employed. Sampling of Mohallahs, the of households, and then adults. Primary, secondary, final.

7. Observation Unit

- Unit of data collection from which information is collected.
- Unit of observation and unit of analysis can be same or different. [Interview head of household (UoO) and collect information about every member (UoA)]

8. *Parameter*

- Summary description of a given variable in population (Mean income of families in the city, mean age)
- Survey research involves the estimation of population parameters.

9. *Statistic*

- Summary description of a variable in survey sample. Mean income/age of the sample.
- Use it for estimation of population parameters

10. Sampling Error

- Probability samples seldom provide statistics exactly equal to parameters.
- Estimation of error to be expected for a given sample.

NONPROBABILITY AND PROBABILITYSAMPLING

Major Alternatives to Sampling

- Probability
- Non-probability

Probability

- Every element in the population has a *known nonzero probability* of selection.
- Used when representativeness is crucial for wider gerenalizability.
- Can measure sampling error.

Non-Probability

- Probability of any element being chosen is unknown. Arbitrary – personal judgment.
- Time and other factors more crucial than generalizability.
- Cannot measure sampling error.
- Projecting data beyond sample is statistically inappropriate.

Types of Non-Probability Sampling

Convenience Sampling

- Also called *haphazard or accidental or grab sampling.*
- Obtaining units who are most conveniently available.
- Person-on-the street interviewed by TV people. Personal judgment.
- May not be representative. Least reliable.
- Used during exploratory phase

Purposive Sampling

- Researcher lays down the criteria for the subjects.
- Also called as *judgmental or expert opinion sample*. Selection depends upon the judgment of selector.

Quota Sampling

- Procedure that ensures certain characteristics of a population sample will be represented.
- Fix the quota. Use convenience sampling. Fieldworker bias in selection.

- Can be considered as proportionate stratified sampling – but on convenience basis
- Speed of data collection, lower cost, and convenience.
- Ensures the inclusion of a subset when it is under-represented.

Snowball Sampling

- Also called: *network, chain referral, or reputational sampling.*
- Analogy of snowball.
- Begins with one/few, spreads out on the basis of links.
- Useful where respondents are difficult to identify.

Sequential Sampling

- Similar to purposive sampling.
- In purposive – get every possible case.
- In sequential – continuous evaluation of data. Stop when no new information is coming.

Theoretical Sampling

- Researcher selects cases based on his own insight.
- Used in observational studies

TYPES OF PROBABILITY SAMPLING

- Requires more work than nonrandom sampling.
- Researcher must identify sampling elements.
- Necessary to contact the sampled respondent. Call back several times.
- Likely to yield representative sample.
- Can determine the relationship between sample and population – Calculate the sampling error i.e. deviation between sample result and population parameter due to random process.

Simple Random Sample

- Assumption – homogeneous population.
- Develop an accurate sampling frame.
- Locate the exact element/s to be selected i.e. decide the sample size.
- Number all the elements in the sampling frame.
- Use the list/table of random numbers.

Table of Random Digits

37751	04998	66038	63480	98442	22245	83538	62351
50915	64152	82981	15796	27102	71635	34470	13608
99142	35021	01032	57907	80545	54112	15150	36856
70720	10033	25191	62358	03784	74377	88150	25567
18460	64947	32958	08752	96366	89092	23597	74308
65763	41133	60950	35372	06782	81451	78764	52645
83769	52570	60950	35372	06782	81451	84990	26400

- Take a random starting point.
- Characteristics of most random samples are close to the characteristics of population.
- Calculate the sampling error.

Systematic Random Sample

- Short cut for simple random sampling.
- Number each element in the sampling frame.
- Instead of using the table of random numbers, researcher develops a system for the selection of elements.

- Calculate the sampling interval. Basis of random selection
- Sample size/Population size X 100 i.e. sample as % of Pop.
- Begin with a random start.
- Draw sample by choosing every Nth case.
- If elements are organized in some pattern in the sampling frame then problem. Possible periodicity in population.

Stratified Random Sampling

- In case population is heterogeneous, then stratify. Each stratum in itself becomes homogeneous.
- Researcher's decision about homogeneity of the population.
- Draw the sample from each stratum by using simple random sampling procedure.

Reasons for stratification

- To increase a sample's statistical efficiency. Each stratum gets represented. Reduces random sampling error. Makes the sample representative.
- Provides adequate data for analyzing the various subpopulations.
- Enables to use different research procedures in different strata.

Stratification Process

- Stratification based on the primary variable (Y). What characteristics of pop affect Y?
- Should increase homogeneity within stratum and heterogeneity between strata.
- Make sampling frame for each stratum.
- Serially number the elements.
- Use table of random numbers and select the sample from each stratum.

Proportionate vs. Disproportionate

- Proportionate: If the number of sampling units drawn from each stratum is in proportion to the relative population size of the stratum.
- Disproportionate: to ensure an adequate No. of sampling units in every stratum. Dictated by analytical considerations

Cluster Sampling

- Purpose: sample economically and retaining probability sampling.
- Heterogeneity within clusters but homogeneity between clusters.
- Random selection of clusters.
- Random selection of elements within the selected clusters.

Cluster sampling addresses two problems

- Researchers lack good sampling frame for dispersed population
- Reaching each sampled element is costly.

Multistage Sampling

- Researcher draws several samples in stages. For example:
- Cluster sampling is usually multistage →

stage1– random sampling of big clusters

stage 2– random sampling of small clusters within the big ones

stage 3– random sampling of elements within the selected small clusters

City blocks – households – individuals

Double Sampling

- When further information is needed from a subset of the group from which some information has already been collected for the same study.
- Want to examine the matter in more detail.
- Sub-sample of the primary sample