Determining the Sample Plan

The Sample Plan is the process followed to select units from the population to be used in the sample

Basic Concepts in Samples and Sampling

- ***Population:*** **the entire group under study as defined by research objectives. Sometimes called the "universe."**

Researchers define populations in *specific terms* such as heads of households, individual person types, families, types of retail outlets, etc. Population geographic location and time of study are also considered.

Basic Concepts in Samples and Sampling

- *Sample:* **a subset of the population that should represent the entire group**
- *Sample unit:* **the basic level of investigation…consumers, store managers, shelf-facings, teens, etc. The research objective should define the sample unit**
- *Census:* **an accounting of the complete population**

Basic Concepts in Samples and Sampling…cont.

- *Sampling error:* **any error that occurs in a survey because a sample is used (random error)**
- *Sample frame:* **a master list of the population (total or partial) from which the sample will be drawn**
- *Sample frame error (SFE):* **the degree to which the sample frame fails to account for all of the defined units in the population (e.g a telephone book listing does not contain unlisted numbers) leading to sampling frame error.**

Basic Concepts in Samples and Sampling…cont.

- ***Calculating sample frame error (SFE):***

  Subtract the number of items on the sampling list from the total number of items in the population. Take this number and divide it by the total population. Multiply this decimal by 100 to convert to percent (SFE must be expressed in %)

If the SFE was 40% this would mean that 40% of the population was not in the sampling frame

Reasons for Taking a Sample

- *Practical considerations such as cost and population size*
- *Inability of researcher to analyze large quantities of data potentially generated by a census*
- *Samples can produce sound results if proper rules are followed for the draw*

Basic Sampling Classifications

- *Probability samples:* **ones in which members of the population have a known chance (probability) of being selected**
- *Non-probability samples:* **instances in which the chances (probability) of selecting members from the population are unknown**

Probability Sampling Methods

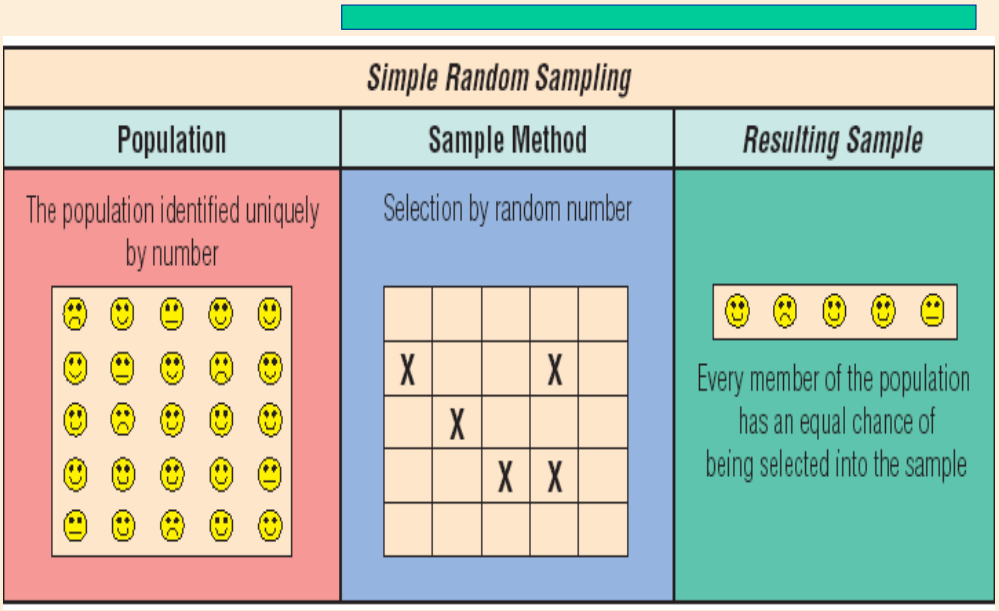
Simple Random Sampling

- *Simple random sampling:* the probability of being selected is "known and equal" for all members of the population
  - *Blind Draw Method (e.g. names "placed in a hat" and then drawn randomly)*
  - *Random Numbers Method (all items in the sampling frame given numbers, numbers then drawn using table or computer program)*
- *Advantages:*
  - Known and equal chance of selection
  - Easy method when there is an electronic database

Probability Sampling Methods

Simple Random Sampling

- *Disadvantages: (Overcome with electronic database)*
  - Complete accounting of population needed
  - Cumbersome to provide unique designations to every population member
  - Very inefficient when applied to skewed population distribution (over- and under-sampling problems) – this is not "overcome with the use of an electronic database)


Simple Random Sampling
Population
Sample Method
Resulting Sample
The population identified uniquely by number
Selection by random number
Every member of the population has an equal chance of being selected into the sample

Probability Sampling Methods

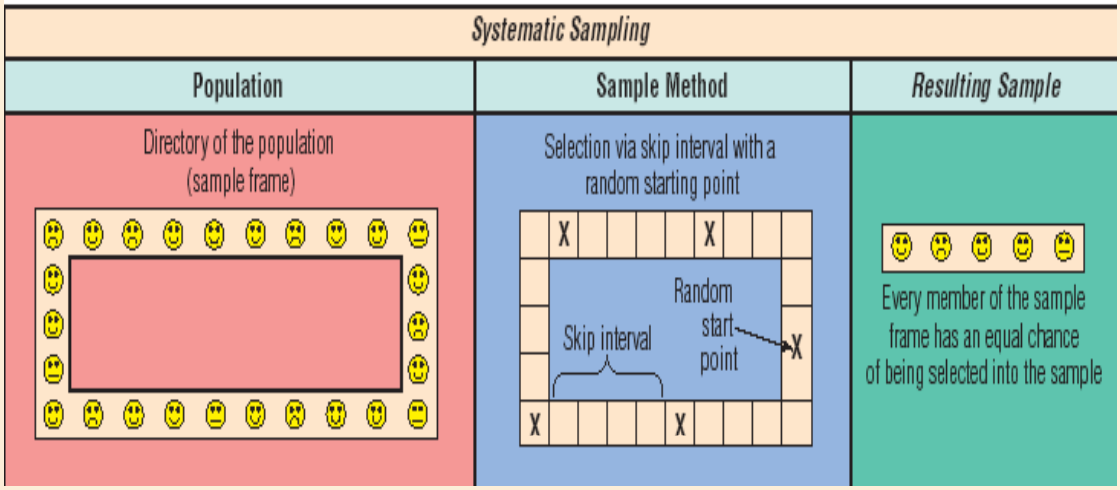
Systematic Sampling (A Cluster Method)

- ***Systematic sampling:*** **way to select a probability-based sample from a directory or list. This method is at times more efficient than simple random sampling. This is a type of cluster sampling method.**
  - ***Sampling interval (SI) = population list size (N) divided by a pre-determined sample size (n)***
  - ***How to draw: 1) calculate SI, 2) select a number between 1 and SI randomly, 3) go to this number as the starting point and the item on the list here is the first in the sample, 4) add SI to the position number of this item and the new position will be the second sampled item, 5) continue this process until desired sample size is reached.***

Probability Sampling Methods

Systematic Sampling

- *Advantages:*
  - Known and equal chance of any of the SI "clusters" being selected
  - Efficiency..do not need to designate (assign a number to) every population member, just those early on on the list (unless there is a very large sampling frame).
  - Less expensive…faster than SRS
- *Disadvantages:*
  - Small loss in sampling precision
  - Potential "periodicity" problems


Systematic Sampling
Population
Sample Method
Resulting Sample
Directory of the population
(sample frame)
Selection via skip interval with a
random starting point
X
X
Random
start
point
Skip interval
X
X
Every member of the sample
frame has an equal chance
of being selected into the sample

Probability Sampling Methods
Cluster Sampling

- ***Cluster sampling:*** **method by which the population is divided into groups (clusters), any of which can be considered a representative sample. These clusters are mini-populations and therefore are heterogeneous. Once clusters are established a random draw is done to select one (or more) clusters to represent the population. Area and systematic sampling (discussed earlier) are two common methods.**
  - ***Area sampling***

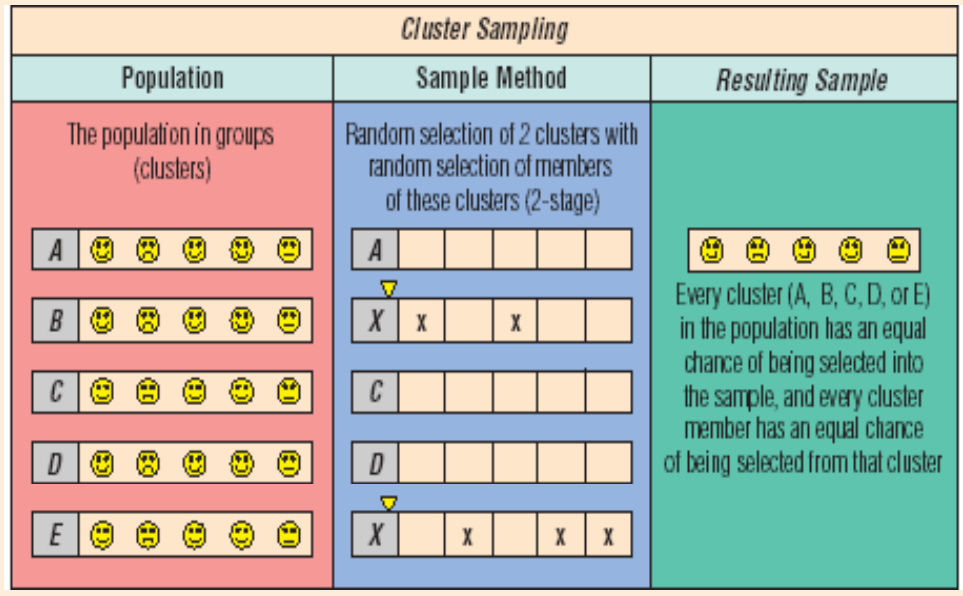
Probability Sampling Methods
Cluster Sampling

- *Advantages*
  - *Economic efficiency … faster and less expensive than SRS*
  - *Does not require a list of all members of the universe*
- *Disadvantage:*
  - **Cluster specification error…the more homogeneous the cluster chosen, the more imprecise the sample results**

Probability Sampling Methods

Cluster Sampling – Area Method

- ***Drawing the area sample:***
  - **Divide the geo area into sectors (subareas) and give them names/numbers, determine how many sectors are to be sampled (typically a judgment call), randomly select these subareas. Do either a census or a systematic draw within each area.**
  - **To determine the total geo area estimate add the counts in the subareas together and multiply this number by the ratio of the total number of subareas divided by number of subareas.**


Cluster Sampling
Population
Sample Method
Resulting Sample
The population in groups (clusters)
Random selection of 2 clusters with random selection of members of these clusters (2-stage)
A
B
C
D
E
X
Every cluster (A, B, C, D, or E) in the population has an equal chance of being selected into the sample, and every cluster member has an equal chance of being selected from that cluster

A two-step area cluster sample (sampling several clusters) is preferable to a one-step (selecting only one cluster) sample unless the clusters are homogeneous

Probability Sampling Methods
Stratified Sampling Method

- This method is used when the population distribution of items is skewed. It allows us to draw a more representative sample. Hence if there are more of certain type of item in the population the sample has more of this type and if there are fewer of another type, there are fewer in the sample.

Probability Sampling Methods
Stratified Sampling

- ***Stratified sampling:*** **the population is separated into homogeneous groups/segments/strata and a sample is taken from each. The results are then combined to get the picture of the total population.**
- **Sample stratum size determination**
  - ***Proportional method (stratum share of total sample is stratum share of total population)***
  - ***Disproportionate method (variances among strata affect sample size for each stratum)***

Probability Sampling Methods
Stratified Sampling

- *Advantage:*
  - More accurate overall sample of skewed population…see next slide for WHY
- *Disadvantage:*
  - More complex sampling plan requiring different sample sizes for each stratum

Why is Stratified Sampling more accurate when there are skewed populations?

The less the variance in a group, the smaller the sample size it takes to produce a precise answer.

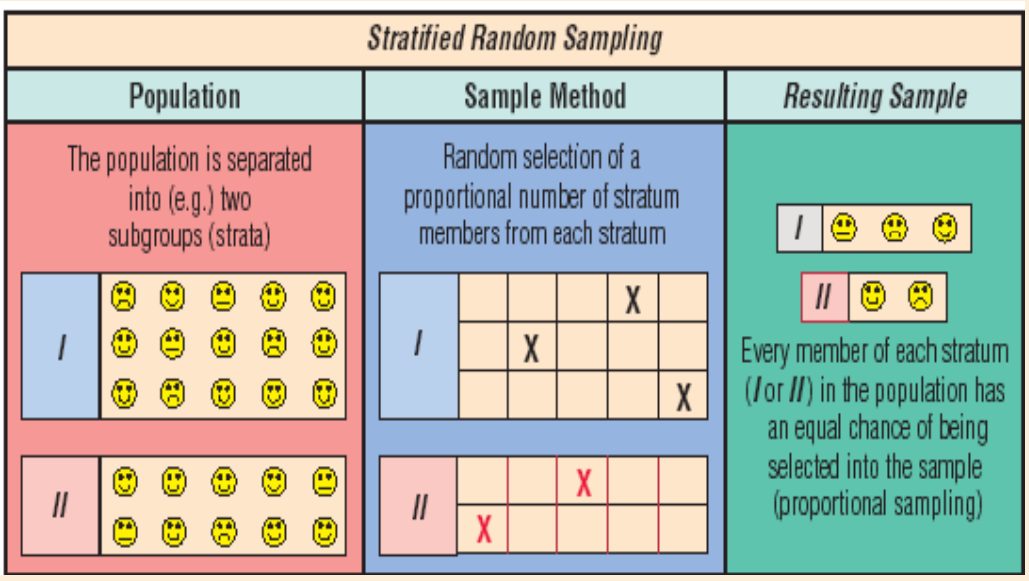
Why? If 99% of the population (low variance) agreed on the choice of brand A, it would be easy to make a precise estimate that the population preferred brand A even with a small sample size.

But, if 33% chose brand A, and 23% chose B, and so on (high variance) it would be difficult to make a precise estimate of the population's preferred brand…it would take a larger sample size….

Why is Stratified Sampling more accurate when there are skewed populations? Continued..

Stratified sampling allows the researcher to allocate a larger sample size to strata with more variance and smaller sample size to strata with less variance. Thus, for the same sample size, more precision is achieved.

This is normally accomplished by disproportionate sampling.


Stratified Random Sampling
Population
Sample Method
Resulting Sample
The population is separated into (e.g.) two subgroups (strata)
Random selection of a proportional number of stratum members from each stratum
I
II
Every member of each stratum (I or II) in the population has an equal chance of being selected into the sample (proportional sampling)

Nonprobability Sampling Methods
Convenience Sampling Method

- *Convenience samples:* **samples drawn at the convenience of the interviewer. People tend to make the selection at familiar locations and to choose respondents who are like themselves.**
  - **Error occurs 1) in the form of members of the population who are infrequent or nonusers of that location and 2) who are not typical in the population**

Nonprobability Sampling Methods
Judgment Sampling Method

- *Judgment samples:* **samples that require a judgment or an "educated guess" on the part of the interviewer as to who should represent the population. Also, "judges" (informed individuals) may be asked to suggest who should be in the sample.**
  - **Subjectivity enters in here, and certain members of the population will have a smaller or no chance of selection compared to others**

Nonprobabilty Sampling Methods
Referral and Quota Sampling Methods

- *Referral samples (snowball samples):* **samples which require respondents to provide the names of additional respondents**
  - **Members of the population who are less known, disliked, or whose opinions conflict with the respondent have a low probability of being selected.**
- *Quota samples:* **samples that set a specific number of certain types of individuals to be interviewed**
  - **Often used to ensure that convenience samples will have desired proportion of different respondent classes**

Online Sampling Techniques

- *Random online intercept sampling:* **relies on a random selection of Web site visitors**
- *Invitation online sampling:* **is when potential respondents are alerted that they may fill out a questionnaire that is hosted at a specific Web site**
- *Online panel sampling:* **refers to consumer or other respondent panels that are set up by marketing research companies for the explicit purpose of conducting online surveys with representative samples**

Developing a Sample Plan

- *Sample plan:* **definite sequence of steps that the researcher goes through in order to draw and ultimately arrive at the final sample**

Developing a Sample Plan
Six steps

- *Step 1:* **Define the relevant population.**
  - ***Specify the descriptors, geographic locations, and time for the sampling units.***
- *Step 2:* **Obtain a population list, if possible; may only be some type of sample frame**
  - ***List brokers, government units, customer lists, competitors' lists, association lists, directories, etc.***

Developing a Sample Plan
Six steps

- ***Step 2 (concluded):***
  - ***Incidence rate (occurrence of certain types in the population, the lower the incidence the larger the required list needed to draw sample from)***

Developing a Sample Plan
Six steps …continued

- ***Step 3:*** **Design the sample method (size and method).**
  - ***Determine specific sampling method to be used. All necessary steps must be specified (sample frame, n, … recontacts, and replacements)***
- ***Step 4:*** **Draw the sample.**
  - ***Select the sample unit and gain the information***

Developing a Sample Plan
Six steps…concluded

- *Step 4 (Continued):*
  - *Drop-down substitution*
  - *Oversampling*
  - *Resampling*
- *Step 5:* **Assess the sample.**
  - *Sample validation – compare sample profile with population profile; check non-responders*
- *Step 6:* **Resample if necessary.**