

FAM SAMPLING WORKSHOP
WESTMINSTER CATERING AND CONFERENCE CENTER
WESTMINSTER, MARYLAND

MONDAY, MAY 24, 1999

02:00 – 03:00

Sampling Frames

- Selection of clusters and households

Sample Selection Procedures

Two-Stage Cluster Sampling

- Most widely used type of cluster sampling
- Definition – process of selecting a sample of elements of interest for a particular survey by first choosing a sample of groups or clusters of elements often referred to as Primary Sampling Units (PSU's) at the first stage of sample selection, and then choosing a sample of elements within sample “clusters” or PSU's
- Why used? – because a list or sampling frame of elements of interest is not available and is generally prohibitively expensive to create (e.g., don't have list of households in project area, but have list of villages)

- Note: cluster samples can have varying numbers of stages of selection:
 - One-stage cluster sample (e.g., take sample villages, take all households in each sample village)
 - Two-stage cluster sample (e.g., take sample villages, take sample of households in each sample village)
 - Three-stage cluster sample (e.g., take sample of districts at first stage, take sample of villages from each sample district at second stage, and take sample of households in each sample village at third stage)
 - More stages of sample selection possible depending upon what is available in the way of a sampling frame
- Number of stages depends upon the type of sampling frame available and whom it is desired to collect data from/about

Sampling Frames

- Definition – a list of sampling units, preferably of the elements of interest for a survey, but more often of aggregations of elements (e.g., villages, city blocks, census enumeration areas) that are used as PSU's
- Sampling frames are a key ingredient in probability sampling – how sampling is undertaken in any particular survey is determined to a large extent by what is available in the way of a sampling frame
- Desirable to have a measure of size for PSU's
- Note: sampling frame must cover entire intended survey universe – if not, danger of bias
- If a control area is to be used, need a sampling frame for the control area as well

Steps in selecting a cluster survey of households

- Choose a sample of primary sampling units (PSU's)
- Choose a sample of households within each sample PSU
- Choose a sample of individuals within each sample household (optional)

Step 1 – Select a sample clusters or PSU's

- Typically done using simple random or systematic sampling
- Two options available:
 - Probability proportional to size (PPS) – each PSU is given a chance of selection that varies depending upon its size
 - Equal probability – each PSU is given an equal chance of selection
- Why use PPS? – to control bias that may result from the fact that PSU's vary in size

Steps in choosing sample clusters PPS and EP & examples – Sampling Guide, Figures 4-1 to 4-4.

Step 2 – select a sample of hh within PSU’s

- Simple random/systematic sampling following listing update (used in DHS and most national surveys)
 - Advantage – most accurate/least prone to bias
 - Disadvantage – most costly & time consuming
- Modified cluster sampling method (see Figures 4-5 & 4-6 in Sampling Guide)
 - Advantages – middle ground between optimal and “quick-and-dirty” methods; probability sample; produces unbiased estimates
 - Disadvantages – requires mapping; more costly & time consuming than random walk method
- “Random Walk” method (EPI cluster survey method)
 - Advantages – least costly & time consuming of options
 - Disadvantages – prone to bias; not a probability sample (the way it is usually done)

Step 3 – Select a sample of respondents in sample households (optional)

- Two strategies for sampling within households:
 - Take all eligible respondents
 - Randomly choose one eligible respondent
- Unless infeasible for some reason (e.g., food balance sheets for children), simpler to take all eligible respondents/subjects, although this will compromise some control over ultimate sample size

Probabilities of selection for common designs

See Sampling Guide, Figure 5-1

Sample Selection for Programs with Limited Coverage

- Three options:
 - Use a list frame of program beneficiaries
 - Screen for program beneficiaries in hh survey interviews
 - Restrict the survey universe to sub-areas of project area where program beneficiaries are concentrated