Sampling for Surveys  
May 23, 2016 – May 24, 2016  
Monday, 8:30 am - 5:00 pm; Tuesday, 9:00am – 5:00pm  
Room: I 2 - 250

Objective
The objective of this module is to provide staff with a clear understanding of the key issues involved in selecting samples of households, enterprises, facilities or individuals to be included in various surveys. The module will cover the basics of sampling, both in theory and in practical terms. The module is useful both for those who plan to be involved in any of a variety of sample surveys as well as those who analyze such data. Upon completing the course the participants should be familiar with the major issues and procedures.

Agenda

Monday, May 23, 2016

Opening
9:00-9:30 OPENING REMARKS AND INTRODUCTIONS

An overview of sampling: why sampling is used, what are its main advantages, the power of probability sampling compared with other forms of sampling and the costs of misuse of sampling techniques

Presenter: Kristen Himelein, GPV01

Session I
9:30-11:15 SAMPLING FOR PROPORTIONS AND AVERAGES

Basic concepts of sampling presented as a simulation (Example: an electoral poll in a small island) and then formalized. Population and sample. Truth and estimation: the meaning of sampling error. The distribution of sampling error. How is sampling error related to sample size? How is sampling error related to the size of the population? How is sampling error related to the proportion that is being estimated? Estimating averages (Example: an income survey in a small island). How is sampling error related to the distribution of incomes in the population? The concepts of simple population, sample, sample frame, estimation, sampling error and confidence intervals will be covered.

Presenters: Jonathan Kastelic, DECSM  
            Michael Wild, DECSM

11:15-11:30 Coffee Break
Session II
11:30-12:15 GROUP EXERCISES

Participants simulate various alternative situations in their own computers. Discussion

Presenters: Jonathan Kastelic, DECSM

12:15-1:30 Lunch (not provided)

Session III
1:30-2:30 STRATIFICATION-CONCEPTS AND FORMALIZATION

Stratification (Example: simulation of an electoral poll in an archipelago). The two main reasons for stratification are [a] improving the precision of the overall estimate and [b] obtaining separate estimates for different groups of the population. These objectives are generally contradictory in practice and arbitration is needed. Key concepts are: analytic domains and strata, allocating the sample among strata, proportional and equal allocations and what are sampling weights and why are they needed.

Presenters: Jonathan Kastelic, DECSM
Michael Wild, DECSM

Session IV
2:30-3:30 MULTI-STAGE SAMPLING AND CLUSTERING

Two-stage sampling: rationale and motivation (Example: simulation of an electoral poll with respondents clustered into city blocks). Comparison with a simple random sample: the price to pay is a larger error. Intra-cluster correlation and cluster effects. Relationship between cluster effect, cluster size and intra-cluster correlation.

Presenters: Jonathan Kastelic, DECSM
Michael Wild, DECSM

3:30-3:45 Coffee Break

Session V
3:45-5:00 THE PRACTICE OF SAMPLING FOR HOUSEHOLD SURVEYS

The national census as the basis of the first-stage sample frame. Selecting Primary Sampling Units (PSUs) with Probability Proportional to Size (PPS). Household listing operations for the development of last-stage sample frames: tools, methods and resources. Selection of households from the listings by Systematic Sampling. Common difficulties: segmentation of outsized PSUs; targeting special groups (disabled, migrants, female-headed households, etc.); alternatives to the household listing operation.

Presenters: Jonathan Kastelic, DECSM
Michael Wild, DECSM
Tuesday, May 24, 2016

Session VI
9:00-10:30 SAMPLE FRAMES – NON-TRADITIONAL SAMPLING METHODOLOGIES

Frames may not exist for a variety of reasons or they may be outdated or inadequate for the observation of elusive groups such as migrants. In these cases alternative techniques are needed to construct good sample frames. Quite distinct examples of constructing frames are presented here, including case studies from non-traditional sampling frames.

Presenters: Kristen Himelein, GPV01
Michael Wild, DECSM

Session VII
10:30-11:15 DESIGN EFFECTS

The combined effect of stratification, clustering and weighing – Design effect and rate of homogeneity. Alternative interpretations of deff (and deftp) and why they cannot be ignored. Typical values found in practice. Information and software needed to compute design effects. Implications for the documentation of the design and implementation of the sample, including a case study from recent research.

Presenters: Jonathan Kastelic, DECSM

11:15-11:30 Coffee Break

Session VIII
11:30-12:15 WEIGHTING

The session introduces basic concepts in weighting for complex sample surveys. Issues to be addressed include basic weight calculation, trimming, post-estimation and a brief introduction to the components of panel weights.

Presenter: Michael Wild, DECSM

12:15-1:30 Lunch (not provided)

Session IX
1:30-2:30 SAMPLING ISSUES IN IMPACT EVALUATION (PIES)

The session covers experimental design issues, including power calculations and the importance of not losing sight of selection probabilities, as well as issues related to the construction of panel surveys.

Presenter: Jonathan Kastelic, DECSM
Session X
2:30-3:00 IMPLEMENTATION OF SAMPLING THEORY

This session will present case studies that apply the principles covered in earlier sessions. Sampling methodologies for constructing nationally representative sample will be discussed. This session will also give participants the opportunity to ‘operationalize’ the theory of the previous two days – including using a dataset to determine design effects, construct a sample design, and select the units for a survey.

Presenters: Jonathan Kastelic, DECSM

3:00-3:15 Coffee Break

Session XI
3:15-4:45 IMPLEMENTATION OF SAMPLING THEORY (CON’T.)

4:45-5:00 Wrap Up and Course Evaluations