Pacific Training on Sampling Methods for Producing Core Data Items for Agricultural and Rural Statistics

13-17 August, Suva, Fiji

Module 2: Review of Basics of Sampling Methods
Session 2.1: Terminology, Concepts and Definitions

By
Chris Ryan
Statistician (ESCAP Pacific Office)

UNITED NATIONS siap
Statistical Institute for Asia and the Pacific

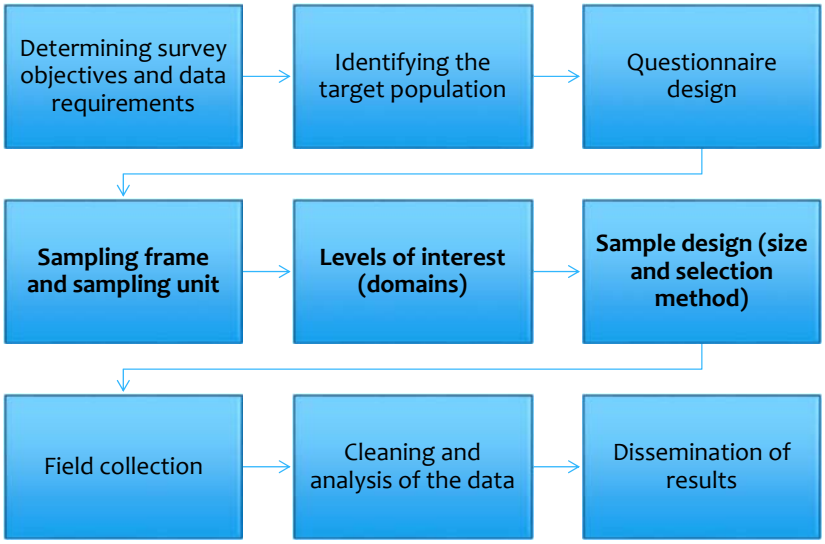
FAO FIAT PANIS

UNITED NATIONS ESCAP
Economic and Social Commission for Asia and the Pacific

Topics Covered

* Steps to Survey Design
* Survey Objectives
* Population of interest
* Target v Sample Population
* Survey Outputs
* Survey Units
* Methods of measurement
* Levels of Interest – domains
* Sample Size

Statistical Survey Design – Steps involved



Survey Objectives

- The starting point of any survey, whether it be a census or sample survey, is to determine the objectives
  - All objectives
  - Not just the main ones
- There are generally numerous objectives associated with a survey
  - Eg, HIES
    - Re-base CPI
    - Provide inputs into National Accounts
    - Poverty Analysis
    - Analyze the importance of Agriculture and Fisheries to Household Income

Survey Objectives (cont)

* How do we achieve this?
  * Need to undertake discussions with key users
* Who are the key users?
  * They're the people pushing for the survey to take place
* The process may involve numerous discussions over a period of time
  * BUT IT'S IMPORTANT TO GET IT RIGHT!!

Survey Objectives (cont)

* It's often the role of the Survey Statistician to assist the key users with determining survey objectives
* Why?
  * Often key users aren't exactly sure what they want or can get from the survey
  * The Survey Statistician is often in a better position to know what is feasible (experience)
  * Costs and Timing are often important factors
    * An area of the Survey Statistician's experience

Target Population v Sample Population

The key first stage of any survey design is knowing the population you wish to make statements about. The population could be anything:

* Small businesses (50 or less employees)
* All households in the urban population of a country
* Households involved in Fishing activities in rural communities

Population of Interest (cont)

* There are different types of populations
  * Target population
  * Sample population
* Target population
  * The set of units we want to collect information from - Scope
* Survey population
  * The set of units we can collect information from – Coverage
* Ideally we want Target population = Survey population

Target Population v Sample Population

Examples of differences

* Remote islands with very small populations not covered
* Homeless people not included in a household Labour Force Survey

Survey Outputs

* Once the Survey Objectives & Population of interest have been determined, the desirable survey outputs should be drafted
* These outputs may be in the form of:
  * Tables
  * Graphs
  * Summary Statistics
* The outputs should relate directly to ALL the objectives of the survey

Survey Outputs

* Outputs should be drafted up to meet all the survey objectives
* If an output is produced which doesn't relate to the survey objectives then either:
  * Add it to the survey objectives if it is required

    or
  * Delete the output

The survey units

* There are different types of survey units:
  * Selection Unit
  * Reporting Unit
  * Tabulation Unit
* Usually these units are the same but not always

The survey units

* Selection Unit: The unit in the sample that is selected
* Reporting Unit: The unit that responds to the survey
* Tabulation Unit: The unit for which results are presented

Example of different units for a survey (HIES)

* Select Dwellings in the survey
* Individuals & Households Report to the survey
* Tabulate Household results

Method of measurement

* The are numerous ways to collect the information
* Some common methods are:
  * Face-to-face interview
  * Self enumeration
  * Mail out survey form
  * Phone interview
  * Administrative records

Method of measurement

- The choice of measurement depends on:
  - The most effective way to get accurate results
  - Cost constraints
  - Time constraints
  - Confidentiality issues
- Many surveys conducted outside the Government (in particular the NSO), ignore the first point and produce unreliable results

Levels of interest (Domains)

- Who do you want to produce results for?
- Is it Agricultural Farmers as a whole, or is it more important to produce results for sub-populations for comparative reasons. For example:
  - By geographical region
    - Urban/Rural, Province, Village, etc
  - By size of farm
  - By type of produce
- The more sub-populations you wish to produce information for the larger the sample size would be.
  - The smaller the sub-population that you have, the larger the sample fraction you are required to select.

Sample Size

Factors which impact on sample size

* Population Size
* Accuracy requirements
* Sample selection methodology
* Domains of interest
* Degree of variation in data being collected

Determining the Sample Size

* In theory, the correct way to determine the sample size for a survey is:
  * Identify the key variables in the survey
  * Determine the accuracy requirements for these variables
  * Select a sample large enough to meet these requirements

NB: the sample selection methodology also impacts on the size of the sample required

Determining the Sample Size

For a simple random sample the formula is:

$$n \geq \frac{1}{\left(\frac{1}{N} + \frac{RSE(\hat{\bar{Y}})^2 \times \hat{\bar{Y}}^2}{s^2} \right)}$$

n = Sample size

N = Population size

$\hat{\bar{Y}}$ = Estimate of mean of variable Y

s^2 = Measure of spread of variable Y

$RSE(\hat{\bar{Y}})$ = Accuracy requirement of variable Y

Determining the Sample Size

- In reality, you generally don't have previous survey data of sufficient quality in which to undertake the necessary calculations
- So what happens?
  - The finances available for the survey generally control the sample size (within reason)
  - You make an educated guess to some degree as to what an appropriate sample size will be