Food and Agriculture Organization
of the United Nations

MASTER SAMPLING FRAME(MSF)

FOR

AGRICULTURAL STATISTICS

Module 3 - Session 1: Sampling design considerations when developing an MSF

2017

Objectives of the presentation

At the end of this session, the audience will:

- Be able to identify the different concepts when designing a survey
- Know the elements that impact the sampling variability and errors
- Understand the particularities of the different probability sampling designs to be considered in deciding which ones are more adapted to the objectives of the survey

Outline

1. Fundamental concepts

2. Sampling variability and probability samples

3. Probability sampling designs

Introduction

- Selecting a useful and efficient sample requires:
  - Knowledge of the sampling theory
  - A Sampling frame designed to meet the country's needs
- Involves a series of steps
  - related to the choice and definition of fundamental concepts in sampling.
- Four questions that should be asked: Who, Where, When and How.

1

Fundamental concepts

1.1. Population or target population

- ***A population or target-population***
  - Finite set of all elementary units about which information is sought

- Three typical elements in agricultural surveys:
  - Holdings or farms
  - Holders or farmers
  - Households or dwellings

- Requires identification of a place and a point in time.
  - Examples of populations:
    - The set of all holders of a province in 2014.
    - The set of all households of a region in a given year.

1.2. Subpopulation

- Multi-purpose aspects of agricultural surveys may require estimates for ***subpopulations of interest***
  - elementary units for which inferences are required.

- Maize production as an item of interest
  - inferences for the subpopulation of holdings with maize production from irrigated lands may be necessary.

1.3. Frame or sampling frame

- ***Sampling Frame*:** Reference system composed by a set of materials, devices or coordinates that allows a sample to be selectedproviding access to their respective sampling units (Ferraz, 2015).

- Three types of frames are commonly cited:
  - Area frames
  - List frames
  - Multiple frames

- Two types of frames according to the level of access to population elements
  A) Frames that identify and give direct access to population elements
  B) Frames that provide indirect access to population elements

1.3. Frame or sampling frame

- Characteristics of a good frame:
  - Provides full coverage of the population of elements [sampled population]
  - Enables the identification and accessibility for each of them

- Set of all farmers as the target population
  - A frame listing each farmer
  - Addresses that provides direct access to each item of interest.

1.3. Frame or sampling frame

- Area frame for this same population
  - segment of area or point would be the sample unit
  - provide direct access to a land parcel
  - capability to measure the items of interest by observation or indirect type of access by interviewing the operator of the parcel in the segment

- For items of interest such as income
  - each selected segment of area must be linked to a farmer
  - often feasible only after a visit to the field.

1.4. Sampled population

- ***Sampled population (also called survey population or frame population)***: population actually covered by the survey

- If the frame is complete, unique and up-to-date

  sampled population = target population

- Frame limitations, however, may lead to discrepancies between the sampled and target populations.

1.4. Sampled population

- This may happen for a variety of reasons:
  - The best available list frame is incapable of covering the population
  - The accuracy of the frame containing this population may change over time

- Rules of association must be established to determine
  - when new households/holdings should be substituted for those no longer in existence

- In any case, the inference is valid for the sampled (frame) population

1.5. Variables of interest

- ***Variables of interest***
  - Characteristics that relate to each item of interest
  - Measured for each element of the population

- If Maize is an item of interest
  - area, yield, and production of maize are examples of variables

- If Income is an item of interest
  - income from crops and livestock are examples of variables of interest

1.6. Parameters

- ***Parameters***
  - *Numerical characteristics relating to each item of interest that are aggregated over the population's elements*
  - *Summaries of the values of the variables of interest*

- Examples:
  - Average crop yields
  - Total area cultivated for a specific crop
  - Percentage of farms using a certain type of transportation

1.7. Sampling unit

- ***Sampling units***
  - Element or a set of elements from the target population
  - Identified through the frame
- Can be directly selected by a randomization process, leading to the sample selection.
- A sample unit is associated to each sample stage when multiple stages of sampling
  - At the first stage, sample units are called Primary Sampling Units (PSUs)
  - At the second stage, sample units are called Secondary Sampling Units (SSUs)

1.8. Observation unit

- ***Observation units***
  - *Units on which the measurement procedure is applied*
- Sometimes, sampling and observation units are the same
- In area sampling, observation units can be:
  - segment of land, if objective measurements are to be taken or
  - the holder or holders associated with this segment of land
- In list frame surveys, the observations unit may be:
  - the holding/ holders/ agric. households
  - and/or the parcel of land being operated

1.9. Reporting unit

- ***Reporting units***
  - *Units that report the required data concerning population elements*

- If such data come from direct measurements, the reporting and observation units are the same
  - Example:
    - a farmer is asked for a subjective estimate of his production on a certain type of crop
    - the farmer is the reporting unit that provides information about the farm (observation unit)

1.10. Analytical unit

- ***Analytical unit***
  - Unit on which analyses are related to
- Most of the time, coincides with the observation unit
- Could be an aggregation of observation unit or another unit related to agregation level of the analysis

1.11. Reference period

- ***Reference period***
  - Time period for which statistical results are collected or calculated and to which, as a result, these values refer
  - May be a calendar year (**reference year**), a fiscal year, a semester, a quarter, a month and even a day

- Important to identify the **reference period** of the survey in particular in the survey questionnaire, analysis and metadata

- **Important consideration: seasonality**
  - Certain activities are are specific to certain weeks, months or years

- Conclusions that refer to a specific time frame may not be valid for other time frames.

1.12. Analysis plan

- ***Analysis plan***
  - how much detail is required for each item and the format of the results?
- Contains detailed plan of the way the data are to be analysed and presented
- Required the availability of a tabulation plan
- Greatly facilitates the design of the questionnaire.

1.13. Examples

- **The Gambia's National Agricultural Sample Survey**

- National-level survey from which the following concepts can be identified: http://www.gbos.gov.gm/nada/index.php/catalog/6#page=overview&tab=study-desc.

1.13. Examples

The Gambia's National Agricultural Sample Survey

- **Target population**: The set of all households in the country engaged in growing crops and/or breeding and raising livestock in private or in partnership with others, for a given period or point in time.

- **Subpopulation of interest**: Given the population description, an example of subpopulation of interest for Gambia's survey could be the set of livestock producers.

1.13. Examples

The Gambia's National Agricultural Sample Survey

- **Frame**:
  - list of EAs available from the last population census
  - once an EA was selected, a list of household clusters (called *dabadas*) was built
  - the households that were mainly agricultural were identified to enable the selection of agricultural *dabadas*
  - all households of each agricultural *dabada* selected were listed to enable selection of the household sample

1.13. Examples

The Gambia's National Agricultural Sample Survey

- **Sampled population**:
  - sampled population is consistent with the target population
  - the list of EAs from the last census is still up-to-date
  - The list of households on each selected *dabada* are also up-to-date.
- **Variables of interest**:
  - a series of questions asked to each householder
  - total number of cattle that are less than one year old, area of maize planted in a specific year, yield, production, etc.

1.13. Examples

The Gambia's National Agricultural Sample Survey

- **Parameters of interest:**
  - Estimates of these parameters obtained by applying estimators based on the sample design to the quantity of grain (in Kg) and to the area of maize fields planted in a given year, respectively.
- **Sampling unit:**
  - primary sampling: EA;
  - the Sub-Sampling Unit: *dabada*
  - final sampling unit: household.
- **Reporting unit:**
  - The holding and all activities associated with it is the reporting unit.

1.13. Examples

The US Agricultural Resource Management Survey

- National-level survey from which the following concepts can be identified:

http://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Chemical_Use/ChemUse FieldCropsStatisticalMethodology.pdf.

1.13. Examples

The US Agricultural Resource Management Survey

- **Target population**:
  - "all establishments that sold or would normally have sold at least $1,000 of agricultural products during the year
  - excluding abnormal or institutional farms" for a reference year
- **Subpopulation of interest:**
  - an example of subpopulation of interest is the set of establishments from the population that received government subsidies
- **Frame:**
  - list frame of farms from the USDA-NASS, accounting for 90 percent of the country's land in farms

1.13. Examples

The US Agricultural Resource Management Survey

- **Sampled population**:
  - set of farms described in the definition of the target population that is included in the 90 percent of farmland covered by the frame.
  - the remaining 10 percent is supposed to have a negligible impact on estimates
- **Variables of interest (example)**:
  - percent of total planted acres that received one or more applications of a specific fertilizer nutrient or pesticide active ingredient

1.13. Examples

The US Agricultural Resource Management Survey

- **Parameters of interest**:
  - correspond to taking an average of the variables of interest mentioned above
- **Sampling unit**:
  - the name of the farm operator or the name of the enterprise
- **Reporting unit**:
  - land operated by the selected name and all agricultural activities associated with that holding

2

Sampling variability and probability samples

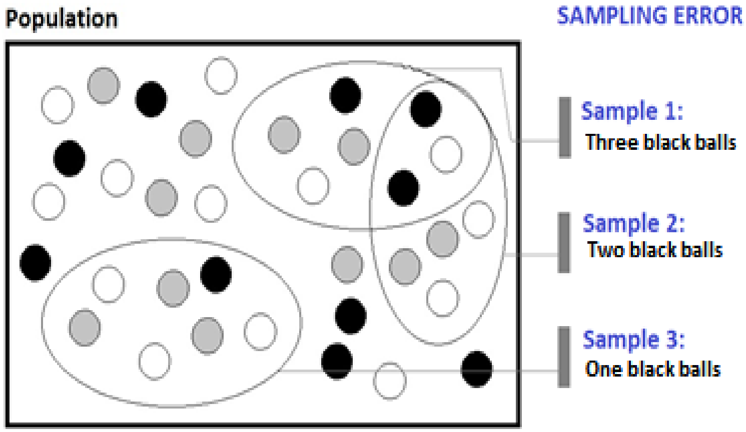
2.1. Sampling variability and probability samples

Two main groups of sources of error in sample surveys:

- **sampling errors**
- **non-sampling errors**

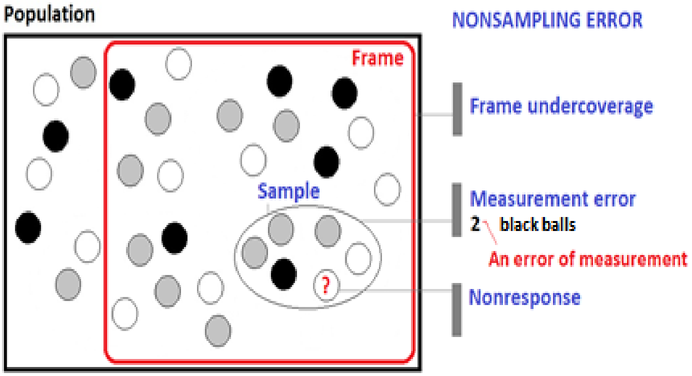
Sampling error:

related to the fact that the variable of interest is observed only for elements in a sample.



2.2. Sampling variability and probability samples

Non-sampling error: related to the general operational process involved in a sample survey.



TOTAL ERROR = Sampling error + Non-sampling error

2.3. Sampling variability and probability samples

Biemer and Lyberg (2003) divided the non-sampling errors in the following categories:

a) Specification errors;

b) Frame coverage errors;

c) Processing errors;

d) Measuring errors;

e) Non-observation errors

- Example of frame error:
  - Imperfections of sample frames
    - lack of complete coverage
    - over-coverage due to duplication
    - poor connections between the target and survey populations

2.4. Sampling variability and probability samples

- Suppose that N is the size of the target population, and let U be the set of indices uniquely identified: $U = \{1,2,\dots,N\}$. Let $S \subset U$ be a sample of n from U.

- Let y_k be the value of the variable of interest y for unit k of the target population U.

- The inclusion of k in the sample is indicated by the following random variable:

$$I_k = I_k(S) = \begin{cases} 1, if\ k \in S \\ 0, otherwise \end{cases}$$

2.5. Sampling variability and probability samples

- Probability sampling designs introduced in the next part determine the exact distribution of I_k, providing the sample inclusion probabilities:

$$\pi_k = P(I_k = 1); \pi_{kl} = P(I_k I_l = 1)$$

- Horvitz and Thompson (1952) as a unifying result
  - ensures the unbiased estimation of parameters such as means, totals and percentages.

2.6. Sampling variability and probability samples

- Consider, from now on, the problem of estimating a population total:

$$Y = \sum_{k \in U} y_k$$

- Thus, the Horvitz-Thompson estimator is given by

$$\hat{Y} = \sum_{k \in S} \frac{y_k}{\pi_k}$$

2.7. Sampling variability and probability samples

- It can be seen that the estimator cannot be applied to nonprobability samples, where the values of π_k are not known and some are zero.
- A general form for the Horvitz-Thompson variance estimator can be written as

$$Var_p(\hat{Y}) = \sum_{k \in U} \sum_{l \in U} (\pi_{kl} - \pi_k \pi_l) \frac{y_k}{\pi_k} \frac{y_l}{\pi_l}$$

- In addition, an unbiased estimate of the variance may be obtained using

$$\widehat{Var}_p(\hat{Y}) = \sum_{k \in S} \sum_{l \in S} \frac{(\pi_{kl} - \pi_k \pi_l)}{\pi_{kl}} \frac{y_k}{\pi_k} \frac{y_l}{\pi_l}$$

2.8. Sampling variability and probability samples

- Requirement for an unbiased variance estimation based on the Horvitz-Thompson approach:
  - probability sample (with $\pi_k > 0$ for every *k* in the population)
  - $\pi_{kl} > 0$ for every *k* and *l*

2.9. Sampling variability and probability samples

- Several authors have addressed the subject of probability sampling. Relevant examples follow:
  - Cochran, W.G. 1977. *Sampling Techniques*. 3rd ed., John Wiley and Sons.
  - Lohr, S. 2009. *Sampling: Design and Analysis*; 2nd edition, Duxbrury Press.
  - Barnett, V. 2002. *Sample Survey: Principles and Methods*. John Wiley and Sons.

3

Probability sample designs

3.1. Simple random sampling

- The probability of selecting a given sample *s* is

$$\boldsymbol{P(s) = \binom{N}{n}^{-1}}$$

- In a simple random sample, the inclusion probabilities are

$$\boldsymbol{\pi_k = \frac{n}{N}} \quad and \quad \boldsymbol{\pi_{kl} = \frac{n(n-1)}{N(N-1)}}$$

3.1. Simple random sampling (SRS)

- Guarantees that each element identified by the frame has the same probability of being included in the sample.

- In case that the selection process is biaised (not really random) the SRS **<u>does not guarantee</u>**:
  - The geographic spread of sampled points.

  Example: all farms selected may be in one corner of the state or country

3.2. Systematic sampling

1) Calculate the step

$$k = \frac{N}{n}$$

Suppose that step (*k)* is an integer number.

2) The first sampling unit is randomly selected from the first *k* elements identified by the frame: let assume it is "b"

3) From the previous unit selected, every *k*-th element of the frame is also included in the sample: *b+k*, *b+2k,…*, *b+(n-1)k*

3.2. Systematic sampling

- If the step (*k*) is not an integer, then a fractional interval method can be carried out as explained below:
- **Step 1**: Randomly select a real number *"b"* from a uniform distribution in the range of $(0,k]$;
- **Step 2**: The first element of the sample is the small integer that is greater than or equal to *b*;
- **Step 3**: choose the integer that is most close to $b+k$, $b+2k$,…, $b+(n-1)k$, based on the calculation below:
  - Let assum the integer "n" such that $n \leq b+k \leq n+1$
  - $d1= |b+k-n|$
  - $d2=|b+k-n-1|$

3.2. Systematic sampling: example

- Example: to select a systematic sample of size n=5 from a population of size N=21, we calculate k=4.2 and proceed to the following steps:

- **Step 1**: Randomly select a real number greater than 0 and smaller than or equal to 4.2. Suppose that this number was b=2.1.

- **Step 2**: The smallest integer greater than or equal to b is 3. Element 3 is the first element of the sample.

- **Step 3**: The next elements will be the smallest integers greater than or equal to $b+k$=6.3, $b+2k$=10.5, $b+3k$=14.7 and $b+4k$=18.9. Therefore, the corresponding elements in the sample are: 7, 11, 15 and 19.

3.2. Systematic sampling

- In systematic sampling, the inclusion probabilities are

$$\pi_k = \frac{n}{N} \quad and \quad \pi_{kl} = \begin{cases} \frac{n}{N}, if\ k\ and\ l\ are\ in\ the\ sample \\ 0, \quad\quad otherwise \end{cases}$$

3.2. Systematic sampling

Advantages

- Ensure geographical spread
- Very straightforward implementation
- Possible to select a sample in the field, if necessary (during survey).

- Example: If villages are selected at the first stage and a listing of farms in the sampled villages is prepared:
  - A simple systematic sample could be collected on the spot

3.2. Systematic sampling

Disadvantages

- Systematic sampling selects a sample from a very limited number of possibilities.
- Although it is a probability sample, it leads to an assignment of probabilities of zero to some π_{kl}.
- Not possible to provide an unbiased estimate for the variance based on the Horvitz-Thompson variance estimator $\widehat{Var}_p(\hat{Y})$

3.2. Systematic sampling

- Replicated sampling design to overcome variance estimation limitations

- Systematic sampling improves the traceability of the sampling process.

- It also makes it easier to prove that the sampling process has not been manipulated.

3.3. Replicated sampling

- The sample procedure involves selecting several small samples, instead of a single large one.

- Example: To select a sample of 10 units from a population of 50:
  - It is possible to select two samples of 5 units
  - The two samples can be selected by simple random sampling or by systematic sampling

3.3. Replicated sampling

- Reasons for using replicated sampling
  - retain the advantage of systematic sampling but to allow sufficient randomization to estimate sampling errors correctly
  - make it easier to rotate samples and make adjustments in sample allocations, when necessary
  - Facilitate the estimation of variance in the context of complex survey design (bootstrap, jacknife etc.)

3.4. Stratified sampling

- Stratified sampling
  - Population divided into subgroups called strata
  - Elements are sampled from each stratum (subgroup) on the basis of a given probability sample design
  - Requires some information on the sample units
- Used when estimates are to be made for subpopulations of interest:
  - Geographic or administrative areas
  - Rare items etc.
- Also used when the size of sample units varies considerably

3.4. Stratified sampling

- Often necessary to define the number of strata
  - Generally, only a few (four or five)
  - More strata will be necessary if the stratification is performed to separate sample units by relative size and by type **(for example, the number of cattle by meat or milk)**
- The table below illustrates the relative efficiency of stratified sampling compared to simple random sampling

3.4. Stratified sampling

Effect of correlation between item and measure of size, and the number of strata, on the relative efficiency of stratified sampling

Number of strata	Correlation between survey item and measure of size				
	.20	.40	.60	.80	.90
	Variance ratio between stratified and simple random sampling				
2	.85	.70	.55	.40	.32
4	.81	.63	.43	.25	.16
6	.80	.61	.42	.22	.11
8	.80	.60	.41	.21	.11

3.4. Stratified sampling

- **If the stratification is performed for geographical or type-of-farm:**
  - The desired breakdown will determine the boundaries

- **If stratification is by size, some general rules of thumb are:**
  - The total of the item being estimated across the strata should be equalized
  - The means should be as different as possible between strata
  - Large, unusual farms or those producing rare items can be placed in separate strata
  - Some strata can be called “pre-select” or “take-all” that is, to be included with certainty if they contain units so large that they would excessively influence the variance (probability of selection=1)

3.5. PPS sampling

- In the previous examples, each population unit had **the same chance of being selected** (in the whole population or the strata)

- **In Probability-proportional-to-size (PPS) sample**
  - A measure of size can be attached to each unit
  - The probability of selection of the unit will depend on its measure of size

3.5. PPS sampling

- The following example is used to illustrate:

Name	Measure of Size	Accumulated Measure
1	10	10
2	1	11
3	4	15
4	15	30
5	5	35

3.5. PPS sampling

- A PPS sample can be selected using either simple random, systematic or replicated sampling
- For example, if a simple random sample of 2 is to be selected, two random numbers between 1 and 35 will be chosen.
- Any random number between 1 and 10 will select Name 1. Only random number 11 will select Name 2.
- To make sure that two unique names are drawn, random numbers are selected until two unique names have been selected.

3.5. PPS sampling

- The inclusion probabilities under an simple PPS design is given by

$$\boldsymbol{\pi_k = min\left\{1, n\frac{x_k}{\sum x_j}, j = 1, 2 \ldots N\right\}}$$

where $\boldsymbol{x_k}$ is the measurement size of unit $\boldsymbol{k}$

$$\boldsymbol{\pi_{kl} = \frac{n(n-1)}{N(N-1)}}$$

3.5. PPS sampling

- Procedures for systematic and replicated sampling can also be used to select samples proportionate to the measure of size.
- To select a sample of two using systematic sampling, first determine the interval 35/2 = 17.5. Then, select a random number between 1.0 and 17.5.
- Again, any random number between 1.0 and 10.0 will select the first sample unit. Next, add the interval to the first random number to determine the second sample unit.

3.5. PPS sampling

- The measure of size comes from a single variable.
  - **For example, in area sampling, the area of a segment can be used as a measure of size to select a PPS sample.**

- The efficiency of PPS sampling lies in the level of correlation between the variable of interest and the variable used as a measurement of size.

3.6. Multivariate probability-proportional-to-size (MPPS)

- Several variables of interest
  - A single measure of size will not properly represent all items
- MPPS:
  - If a series of potential size measures can be identified in the frame, all of them are used in the process of generating an improved size measurement
- The accumulated measures of size can be used as the selection variable

3.6. Multivariate probability-proportional-to-size

- Suppose that there are $J \geq 2$ variables of interest (items)

- Each having at least one auxiliary variable that can be used as a measurement of size.

- Let x_{jk} be the value of the size measure *j* for element *k* in a given *f* frame. Let also

$$X_j = \sum_{k \in f} x_{jk}$$

be the total of the auxiliary variable *j* over frame *f*.

3.6. Multivariate probability-proportional-to-size

- In addition, let n_j be the sample size needed for the variable of interest *j*. Then, the inclusion probability under an MPPS design is given by

$$\boldsymbol{\pi_k^f = min\left\{1, max\left(n_j \frac{x_{jk}}{X_j}, j = 1, 2 \ldots J\right)\right\}}$$

- The remaining steps for selecting the sample are identical to PPS sampling.

3.6. Multivariate probability-proportional-to-size

1. Select a GENERAL sample with n=5 for cropland and n=5 for capacity

Record Number (1)	Control Data Cropland (2)	Control Data Capacity (000) (3)	Relative Measure Col 2 / 10000 (4)	Relative Measure Col 3 / 10000 (5)	Probabilit y of Selection Col 4 *n (6)	Probabilit y of Selection Col 5 *n (7)	Max Prob of Selection Max(6,7) (8)	Random Number (9)	In Sample = 1 (10)	Exp Factor 1 / col 8 (11)
1	2000	1200	0.200	0.120	1.000	0.600	1.000	0.1603	1	1.000
2	1800	800	0.180	0.080	0.900	0.400	0.900	0.9856		
3	1500	700	0.150	0.070	0.750	0.350	0.750	0.2247	1	1.333
4	1400	600	0.140	0.060	0.700	0.300	0.700	0.4889	1	1.429
5	800	700	0.080	0.070	0.400	0.350	0.400	0.0972	1	2.500
6	600	300	0.060	0.030	0.300	0.150	0.300	0.8641		
7	500		0.050	0.000	0.250	0.000	0.250	0.7299		
8	400	500	0.040	0.050	0.200	0.250	0.250	0.9874		
9	300		0.030	0.000	0.150	0.000	0.150	0.1318	1	6.667
10	250	150	0.025	0.015	0.125	0.075	0.125	0.1530		
11	200	100	0.020	0.010	0.100	0.050	0.100	0.2952		
12	100		0.010	0.000	0.050	0.000	0.050	0.3829		
13	60	100	0.006	0.010	0.030	0.050	0.050	0.2283		
14	50	50	0.005	0.005	0.025	0.025	0.025	0.4382		
15	30		0.003	0.000	0.015	0.000	0.015	0.6579		
16	10		0.001	0.000	0.005	0.000	0.005	0.2825		
17		1800	0.000	0.180	0.000	0.900	0.900	0.2366	1	1.111
18		1500	0.000	0.150	0.000	0.750	0.750	0.8459		
19		1000	0.000	0.100	0.000	0.500	0.500	0.0659	1	2.000
20		500	0.000	0.050	0.000	0.250	0.250	0.9685		
Total	10000	10000	1	1	5	5	7.47		7	16.040

Source: USDA

3.7. Cluster sampling

- Main characteristic: **the sampling unit is a cluster of units**
- To select a cluster sample (or single-stage cluster sample), a simple random sample of clusters is taken and each unit in the selected clusters is investigated
- Systematic or replicated sampling can also be used to select a cluster sample
- The inclusion probability in the case of SRS is $\pi_{ik} = \frac{m}{M}$

  with k the first stage (PSU) and *i* the element of the PSU, *m* the number of PSU to be selected, *M* the total number of PSUs

3.7. Cluster sampling

- Cluster sampling for agriculture is a powerful tool to :
  - Optimize travel time for field staff
  - Reduce field costs
  - Develop a Sampling frame
- The efficiency of cluster sampling improves when the variability of the sampling units within clusters is large.
- However, since clusters for agriculture are defined geographically, they tend to be homogeneous.
  - More clusters will have to be selected and then subsampled, using measures of size -> multi-stage sampling

3.8. Two-stage sampling

- Results when an extra stage is added to cluster sampling
- Suppose that 50 farms clustered into 15 villages are to be surveyed.
- Suppose further that it is decided to select five villages at random, obtain a listing of all farms within each selected village, and then select two farms from within each village.
- In this case, each farm has a chance of appearing in the sample at least once with each of the other farms, and the overall sample size and survey workload can thus be controlled.

3.8. Two-stage sampling

- Sampling variability is **usually larger compared to** (single-stage) simple random sampling

- Important!!!!
  - Auxiliary information be obtained to guide the first- and second-stage selections
  - Cost factors must be considered, that is the cost of building a complete frame as opposed to the additional survey costs deriving from a larger sample using two stages of sampling

3.8. Two-stage sampling

- The inclusion probability in the case of simple random selection is:

$$\pi_{ik} = \frac{m}{M} * \frac{n_i}{N_i}$$

With k the first stage (PSU) and i the unit of the PSU , m the number of PSU to be selected, M the total number of PSUs, n_i the number of SSUs to be selected among N_i elements of the PSU.

- The main reason for two-stage sampling is the reduction in survey and frame-building costs

- In two-stage sampling, stratification, PPS or MPPS can also be used to improve the survey design

Conclusion

During this presentation:

- Concepts and sampling considerations have been discussed
  - Will serve as a basis to understand and develop sampling strategies

- The probability sampling designs have been developed:
  - Countries should use the appropriate sampling design
  - Customize regarding to their information available and the budget

References

- Barnett, V. 2002. *Sample Survey: Principles and Methods*. John Wiley and Sons: Chichester, UK.
- Cochran, W.G. 1977. *Sampling Techniques*. 3rd edition. John Wiley and Sons: Chichester, UK.
- Global Strategy to improve agricultural and rural statistics., 2017. *Master sampling frames for agriculture: supplement on selected country experiences.* Rome, Italy
- Global Strategy to improve agricultural and rural statistics., 2015. *Hanbook of Master Sampling Frame for Agricultural Statistics: Frame development, Sample design and Estimation*., Rome, Italy

References

- Horvitz, D.G. & Thompson, D.J. 1962. *A generalization of sampling from a finite universe*. Journal of The American Statistical Society, 47:663-685
- Lohr, S. 2009. *Sampling: Design and Analysis*. 2nd edition. Duxbury Press: Pacific Grove, CA; US.
- Statistics Canada. 2003. *Survey Methods and Practices*. Catalogue n° 12-587-X. Ottawa, Canada
- UNSD. 2005. Designing Household *Survey Samples: Practical Guidelines*. United Nations Publication: New York, US

Thank You