Food and Agriculture Organization
of the United Nations

LIVESTOCK PRODUCTION AND PRODUCTIVITY

Session 3:
Data collection and survey design

2018

Objectives of the presentation

- To present/discuss data collection and survey design

- To develop the capacity of technical staff dealing with survey design and data collection in statistical services in charge of livestock statistics

Outline

- **Introduction**
- **1) Main data sources**
- **2) Data collection methods**
- **3) Survey design and implementation**

Introduction…

- **Variety of sources for livestock data exist**;
  - Agricultural censuses/surveys
  - Other sources e.g. commercial records, administrative records, reporting systems, etc.
- **Some sources are more appropriate/of better quality than others depending on:**
  - Specific variables/indicators
  - Agricultural holding status (whether household or non household sector such as large scale commercial farm enterprises)
  - Livestock rearing system
- **If more than one source available:**
  - Do quality assessment of sources to identify short comings/gaps of each source
  - Identify possibilities of double collecting data and avoid duplication of effort
  - Ensure coherence and additionality when finally compiling from sources

1 Main Data Sources

1.1 Censuses: Agricultural or Livestock

Conducted usually every 10 years, they enumerate all agriculture holdings covering crop, livestock activities

Advantages	**Issues**
Good coverage of permanently settled livestock	Miss nomadic livestock most of the time
Is starting point for sample survey frame building	Expensive
Suitable for herd structure data collection	Long data collection, processing, dissemination

1.2 Periodic and ad-hoc surveys

Use sample (between 1 to 10% of population) to collect regular inter census data

Advantages	Issues
• Cheaper, quicker	At times, the survey of single household is not enough: may need to do communal survey
• Collects data which is detailed/focused	
• Can be annual, bi-annual, quarterly, monthly	

1.2 Periodic and ad-hoc surveys . . .

Methodological considerations to note . . .

Item	Note
Survey objective	To obtain livestock and animal production (at times) at same time data cropping activities of the holdings
Statistical unit	Agricultural holding (see WCA 2020, p. 6.2)
Scope	• Produce estimates of the size & structure of the herd per livestock species • Collect livestock production and utilisation (meat, milk, eggs, wool, honey etc.). • Type and importance of animal diseases, type and quantities of feed used, grazing area used, farm buildings, labour input, etc.

1.2 Periodic and ad-hoc surveys . . .

Methodological considerations to note . . .

Issues	Items to note
Coverage	• entire country (estimate data for excluded parts of country) • cover the entire population of the livestock species
Frequency	• annual, semi-annual, quarterly or monthly • consider data users' needs, human & financial capacity
Reference day & reference period	• may vary according to the survey objectives and scope

1.2 Periodic and ad-hoc surveys . . .

Periodic surveys of livestock input and output industries …

Advantages	Methodology to use
Best for collecting variables: • animal live and carcass weight • availability of agro-industry feed • causes of death • type of veterinary services used	• census or sample surveys
	• administrative registers and accountancy books kept
	• target respondents is slaughterhouses, other slaughtering points, agro-industry, veterinaries

1.2 Periodic and ad-hoc surveys . . .

Household income-expenditure surveys . . .

Advantage	Issues
can be used to collect data on productivity of livestock kept by households	• focus is more on socio-economic data of household
	• excluded are non-household units i.e. enterprises
	• share in the total animal production by enterprises may be large

1.3 Administrative reporting system

- Data collected on a regular basis and at the lowest administrative unit,
  - such as the village or district level

- Sent up the chain to the national statistical agencies
  - process and collect various key livestock data and statistics

1.3 Administrative reporting system . . .

Advantages	Issues
Already existing infrastructure	Duplication of effort in data collection
Livestock extension officers present	Comparability of data with survey results
Good link with livestock keepers & markets	
Can include admin register of livestock - using tracing devices: identification, movement, health & disease data	Admin registers suitable for large livestock not small ones

1.4 Focus group surveys

Administered on groups selected with probability (e.g. those primary sampling units in a probability sampling design)
Questionnaires are completed by eligible respondent farmers (village head, assisted by most knowledgeable persons)

Advantages	**Issues**
helps obtain contextual/background data	If administered through focus group discussions, reliability of the resulting estimates cannot be quantified
Helps obtain data of groups	

1.4 Focus group surveys . . .

- **Example: poultry**
  - Proportions of a poultry producing group engaged with (i) egg production (ii) meat production (iii) both
  - Of the egg producers referred to above, the proportion that (i) sell all their eggs (ii) consume all their eggs (iii) hatch some or all of their eggs.
  - Of the egg producers referred to above that hatch some of their eggs, the proportion that (i) remove eggs from other hen's clutches for establishment of a sitting (ii) allow single hens' clutches to build up (iii) employ some other method.

1.5 Focus groups vs Individual surveys

Consideration	Focus groups	Individual survey
Desired depth of engagement of participants	Strong	Weak
Clarity of questions and responses	Moderate	Strong
Facility to capture individual farm effects	Limited to proportional piling for establishing distribution of individual effects	Strong
Precision of "lists" vs meaning of results	Lists readily available, with little insight into interpretation. Ranking of listed items subject to peer pressure.	Lists difficult to assemble. Ranking from an existing list is achievable.
Flexibility	Strong	Weak
Efficiency of resource utilization	Strong where capacities are fully utilized by large numbers of participants	Can depend on logistics associated with the sampling strategy
Building on-going relationships	Useful	Less useful
Access to gender and minority information/participants	Achievable either by sub-sampling at or before the workshop, or the creation of special sessions at the workshop	Depends on who is available and approachable at the specific time of the survey interview
Control over local partners' action	None	Some
Mapping and other respondents' identity issues	Loss of anonymity amongst peers and neighbors is inevitable Locations and farm-specific information is provided subject to few checks	Anonymity is maintained. Respondent statements about farm size and operations, and other important details, can be checked to some extent
Payment issues	Payments made to participants are usually necessary, at least in the form of transport allowances and meal expenses	Payments are usually not made
Sampling issues	Left in the hands of local contacts, selection for a workshop or communal activity may contribute to a biased sample, or to an inaccurate sample such as one featuring participants that adopt roles related to payments being made	Structured and meaningful sampling is easier to achieve

1.5 Focus groups vs Individual surveys

- Clearly define & identify animal production systems to collect information using communal surveys

- Data categories used should be:
  - Mutually exclusive (i.e. separated by boundaries that are clear to providers and users of information)
  - Exhaustive (i.e. the list represents the full list, or the list of the most frequently expected selections)
  - Manageable in number (i.e. the list is short enough to be usable, a situation usually enabled by a category called "other" and composed of less frequently employed options)

1.6 Technical Conversion Factors (TFC)

TCFs are widely used by the scientific community to convert an easily measured livestock variable into a different unit of measure

Advantages	**Issues**
Reflect well long term trends	Don’t capture short term variations
A good alternative to admin/survey data	Not based on national sample
Can be used in questionnaire piloting/trialling	Require regular reviews/updates to maintain accuracy

1.6 Technical Conversion Factors (TFC)

Examples . . .

Required task	TFC to use
Fill missing item or non-response in the survey e.g. milk production	milk yield per cow per day i.e. multiply number of milking cows estimated from a sample survey
Estimate meat production with only data of number of animals slaughtered	use carcass weight per animal species
Estimate egg production counting number of laying hens & the number of clutches per year	use eggs per laying hen per clutch

2

Data Collection Methods

2.1 Interviews

You can use a face-to-face interviews to collect information on livestock number and animal production

Most often used to collect information on livestock

Tools for Interview	Constraints
Paper and pen data collection (PAPI)	Interview is based on farmer's recall or estimation; reliability
Computer Assisted Personal Interview (CAPI) – on smart phones/tablets	Relies on the respondent's knowledge
Survey software as Survey Solutions, ODK, CSPro	Requires willingness to cooperate

2.2 Reporting

Reporting or self-administrated surveys are based on questionnaires completed by the respondent without the assistance of an interviewer

Suitable for	**Issues**
If respondents keep reliable records	Requires literacy on the part of the respondent
	Requires willingness to reply backed up by legal obligation to report

2.3 Direct observation and measurement

The most accurate method for indicators for which farmers' knowledge or estimation is not reliable. Observation & measurement are done along with questionnaire

Suitable for	Issues
Animal dimensions — weight, heart girth measurement, shoulder height measurement, or body condition score	It is costly if the number of animals to be observed is large
Trained enumerator visually counts the livestock	It is better used on sub sample
Can be used even to revise/improve TCFs	Enumerator needs more training to do this

2.4 Construction of the gold standard

In practice and in context of livestock production and productivity, it means directly measuring a variable and constructing indicators from it, then compare the results to evaluate other methods such as interviews using recall, communal discussions, and various forms of survey question

Approach	**Issues**
1. Construct gold standard	cost of measurement, logistics, personnel, equipment; also inaccurate measurement may occur
2. Test existing method/questionnaire against gold standard also alternative method/ questionnaire against gold standard	hard to define disaggregating variables e.g. types of production system which can be examined together or separately
3. Use gold standard as basis to compare existing vs alternative results in terms of accuracy. Test also cost, logistics etc	because of complexity of undertaking, sample size used may be small weakening statistical inference

3

Survey design and implementation

Introduction

- In the next slides we look at **obtaining data**
- First is a quick reminder of planning the survey
- We examine operational issues in implementing livestock surveys
- We examine data collection specifics according to locality and entity peculiarities
- We finalize by the Sampling Design and Estimators

Planning the survey . . .

- Defining data collection objectives and methods
  - consider country needs
  - consult stakeholders
- Designing the questionnaire
  - can use short, and expanded version depending on needs
  - e.g. FAO-WB-ILRI Livestock Data Innovation in Africa Project also Global Strategy – AGRIS questionnaire - Global Strategy technical assistance questionnaire for Botswana
- Data collection
  - conduct a pilot survey to test – survey design, data collection tools , equipment
  - train enumerators , have in place data collection manual
- Define eligible respondent(s)
  - most knowledgeable person in the holding

3.1 Operational issues in implementing livestock surveys

- **Quality of reporting herd size –** not reporting actual herd size fearing taxes, single herd having animals belonging to more than one owner etc.
  - to avoid **omission** or **duplication,** report information of herd grazing under single management regardless of ownership of the animals
  - livestock belonging to other agric. holding found in other holding temporarily e.g. for sanitary cleaning, should be reported by holding that owns the livestock
  - If two holdings share management of herd, herd should be reported under one holding alone or split between them

3.1 Operational issues in implementing livestock surveys

- **Cross checking quality of herd size information** through:
  - A direct interview with agric. holder on composition of herd (size, breeds, age, sex, destination) by declaration
  - Direct enumeration by the enumerators
  - Veterinary or administrative records on number of vaccinations, births, diseases, slaughtered animals
  - Interview with village herds or communal survey of small group of holders with good knowledge of herd sizes in community
  - Previous studies and research estimates on structure of herds, fertility, mortality rates by breed and locality

Statistical office should examine reliability of above methods and decide on cross checking method(s) before proceeding to actual data collection

3.1 Operational issues in implementing livestock surveys . . . Cont'd

- **Reference day & period, timing of survey and frequency of survey** – three important different concepts

  - **reference day,** refers to the point of time when the number livestock is recorded
    - Use same reference day for all agric. holdings instead of the actual date of the interview to avoid duplication or omission of animals
    - Consider seasonal peaks in production and consumption, events of transhumance when setting reference day as these factors influence number of animals recorded

3.1 Operational issues in implementing livestock surveys . . . Cont'd

- **Reference period** covers a certain time before the reference date and used usually to collect data on livestock production and productivity (e.g. milk produced in last 12 months, eggs laid in last 4 months, no. of slaughtered animals in last month).
  - **Same reference day and reference period** must be used in all surveys over time to allow **time series data** to be built

3.1 Operational issues in implementing livestock surveys . . . Cont'd

- **Reference day & period, timing of survey and frequency of survey** – three important different concepts
  - **Timing of survey** refers to the time selected for the field work.
    - Time the survey to be as close to the reference day as possible
    - Data collection may start on reference day and last from one day to months after reference day (e.g. census).
    - Make period of data collection short (two to three weeks max.) as the quality of data may seriously be reduced when moving away from reference day
    - Consider factors as climate, social and economic conditions etc. since extreme weather, cultural events or periods of heavy agric. work may make respondents less available

3.1 Operational issues in implementing livestock surveys . . . Cont'd

- **Frequency of survey** refers to how often a regular survey is conducted
  - Could be a single collection or repeated collection at set intervals.
  - Frequency of data collection depends on seasonality of production, breeding cycle of the livestock, data importance, user needs and capacity (human & financial) to carry out the survey.
  - Lower respondent recall/too much respondent burden can result from bad frequency and timing of data collection affecting negatively quality of data collected
- Annual survey suffices for cattle herds on farm – births, deaths, sales, purchases, etc.
- Quarterly even monthly data collection for poultry and egg production, pig fattening & slaughtering

3.1 Operational issues in implementing livestock surveys . . . Cont'd

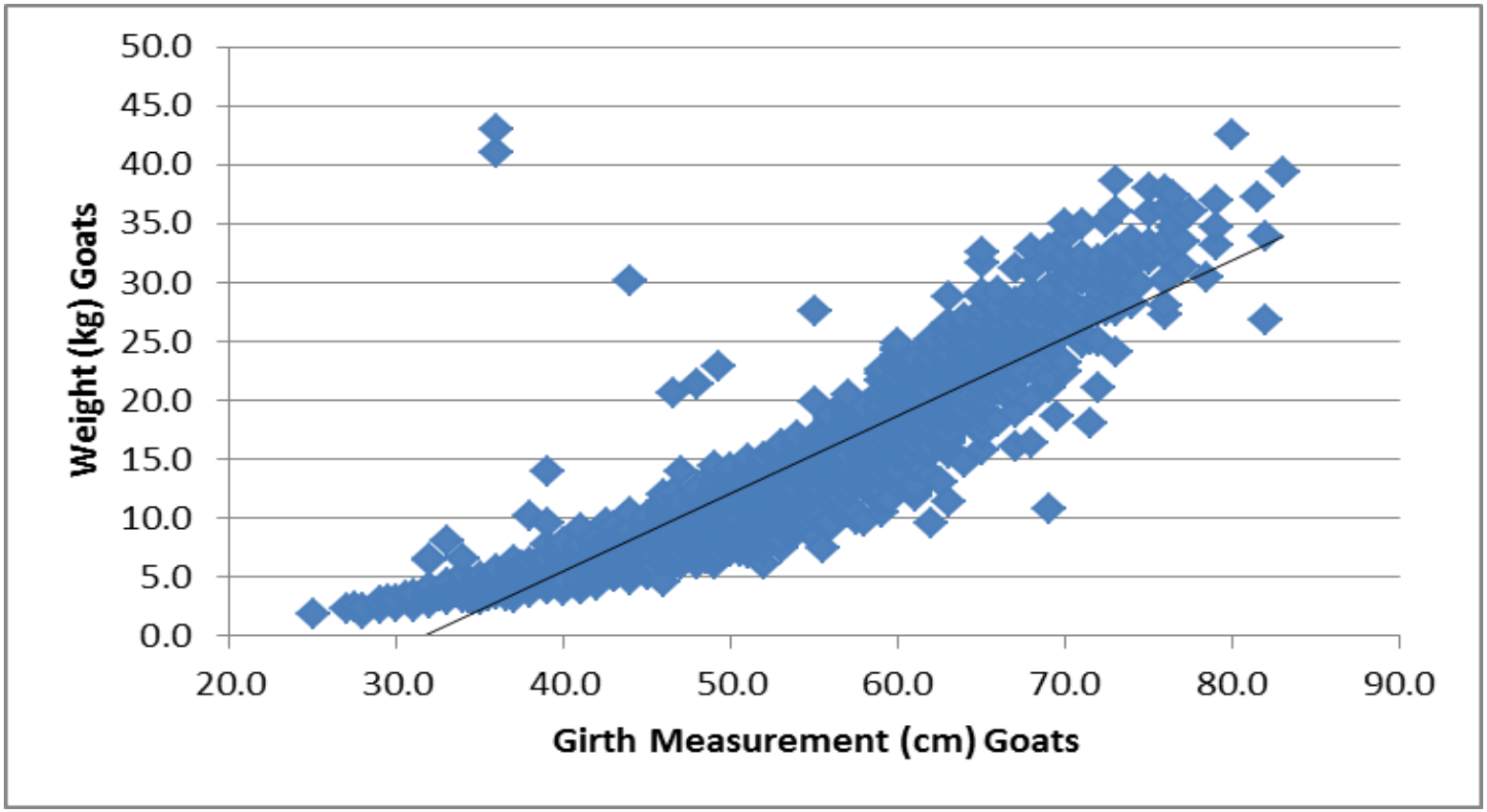
- **Point of survey**
  - refers the location where the data collection takes place
  - depends on location of the holder/respondent & livestock system practiced by the holding (i.e. sedentary, semi nomadism or full nomadism)
  - the respondent leads or shows the enumerator the point of survey

3.1 Operational issues in implementing livestock surveys . . . Cont'd

Estimation of meat production and on farm slaughtering

- **Meat production**
  - Is estimated through growth of live weight of livestock
  - Gold standard measurement through direct weighing of the animals (body condition scores, girth measure, shoulder heights etc)
  - Sample of livestock from herd selected for measurement
  - Enumerator should be able to handle livestock to take different livestock dimensions
  - Alternative get live weight of animals by age, sex, breed at markets
  - Use the table of proxy relationship to estimate meat production. See example

3.1 Operational issues in implementing livestock surveys . . . Cont'd



3.1 Operational issues in implementing livestock surveys . . . Cont'd

Estimation of meat production and on farm slaughtering

- **Meat production**
  - Alternatively use **Off-take** rate
  - **Off-take *rate*** *is the total animals sold, slaughtered or otherwise disposed of (given as gift, donation, etc.) regardless of their future use (meat for home consumption, for sale, share or donation; livestock for slaughtering or for breeding) divided by the total number of animals of the herd*
  - One may cross check there is no under reporting off-take rates for fear of e.g. taxes as it's the case in some countries

3.1 Operational issues in implementing livestock surveys . . . Cont'd

Estimation of milk production

- Estimation based on observed milk production over three days, is good
- Proxy measurement of milk production is still being examined
- Proxy measures tested used cows' Body Condition Score as an indicator of milk production
- Results with Global Strategy depicts statistically significant, although weak, correlation between the proxy measure and milk production.
- Further consideration particular proxy measure is that factors predisposing to good body condition (primarily, good nutrition) are also associated with good milk production

3.1 Operational issues in implementing livestock surveys . . . Cont'd

Estimation of Egg production

- Ask questions on egg production without forgetting to note consumption of eggs by the household
- Ask questions that help effectively identify production systems
- Breed is an important disaggregation variable that helps derive good production and productivity data
  - poultry breeds vary within samples within the same production systems
  - breed is associated with differences in egg productivity
- Meat and egg productivity vary by age and sex structures

3.2 Data collection specifics according to locality and entity peculiarities

- Populations of interest are divided into non-household sector and (small) household holdings
- Two different surveys carried out as result
- Different sources and different data collection methods/tools used
- Harmonize concepts and definitions, reference day and period, list of variables and indicators to ensure that complete and reliable estimates for the entire livestock population are produced

3.2 Data collection specifics according to locality and entity peculiarities

Holdings from the non-household sector (commercial holdings)

- No single international definition
- Each country may have its own definition according to the farm structures and local context
- Generally, these farms are well organised, specialized and market oriented
- May be owned by private or governmental enterprises but also families or households

3.2 Data collection specifics according to locality and entity peculiarities

Holdings from the non-household sector

- usually keep large herds, have permanent employees
- are often registered in business registers, pay taxes and report to statistical office or some other institution on a regular basis
- usually keep detailed records on the number and dynamics of the herds, quantities of production, quantities and value of sales etc.
- Commercial holding own records or administrative data may be enough to obtain livestock production and productivity data

3.2 Data collection specifics according to locality and entity peculiarities

Holdings from the household sector (small holdings)

- raise livestock and usually crops alongside
- may not keep accurate records
- data collection relies of farmer recall
- farmer answers are supported by
  - direct measurement done by the enumerator
  - communal surveys
  - use of technical coefficients
  - data from administrative reports and records

3.2 Data collection specifics according to locality and entity peculiarities

Choose the most appropriate method considering:

- national or regional peculiarities
- sample size requirements
- length of the data collection
- capacity and willingness to participate of the farmers, markets & slaughter houses
- costs of the operation

3.2 Data collection specifics according to locality and entity peculiarities

Better data collection outcomes can be achieved by:

- advocating farmers on importance of livestock data collection thus avoid giving falsified data
- training and involving farmers on data collection
- involving livestock extension officers
- making multiple visits to same sample

3.3 Sampling and estimation techniques

- We will discuss:
  - The definition of target population to measure livestock productivity
  - The main sampling designs for small holdings and commercial holdings
  - The survey preparation, data editing and estimation of population

3.3 Sampling and estimation techniques

- **Target population:** agricultural holdings with livestock raising activities (e.g. raising of cattle, sheep, goats, pigs, poultry … camels if applicable)
  - a threshold on number of livestock/farm income to be taken into account when defining the target population
- Define separate target populations for commercial and small holdings, as different survey designs will likely apply to them
- Consider the type of livestock system practiced by the agricultural holdings to improve the effectiveness of the data collection strategy

3.3 Sampling and estimation techniques

- **We will discuss:**
  - Sample Design (for small holdings & commercial holdings)
  - Frame building (for multistage)
  - Sample size determination & sample allocation
  - Stratification
  - Sample selection
  - Survey weights
  - Estimation (Population total, means and ratios)

3.3 Sampling and estimation techniques

Sample Design for Small Holdings

- Small holdings mostly found in rural areas/types of settlements covering a substantial area of a country
- With limited resources that could be set aside for a livestock survey, a viable design would be multi-stage sampling
- Multi-stage sampling involves sampling from clusters that is known to contribute to higher variability in the survey estimates
- Techniques like stratification and the application of uniform selection probabilities can mitigate the high design effects

3.3 Sampling and estimation techniques

Example of Sample Design for Small Holdings:

- Consider two-stage sampling design
- Given a measure of size B_α for PSU α, we apply a uniform selection probability in selecting PSUs and **b** farming households
- PSUs can be selected using probability proportional to size (PPS) sampling while farming households may be selected using simple random sampling (SRS) or systematic sampling
- The selection equation for selecting farming household β in PSU α will be:

$$P(\alpha\beta) = P(\alpha)P(\beta|\alpha) = \frac{aB_\alpha}{\sum_\alpha B_\alpha} \times \frac{b}{B_\alpha} = \frac{ab}{\sum_\alpha B_\alpha}$$

3.3 Sampling and estimation techniques

- More intervening steps are needed to achieve uniform selection probabilities with PPS
- They include combining small PSUs or deconstructing big ones so that PSU sizes will not vary widely
- The measure of size of PSUs can be derived from the census of agriculture or livestock census
- If not from previous census, data can be obtained from the administrative reporting system
- This measure is number of livestock farming households in a PSU or some other related characteristic correlated to the main variables of interest

3.3 Sampling and estimation techniques

Sample Design for Commercial Holdings

- Commercial holdings are not as predominant as holdings hence no need to apply multi-stage sampling in selecting them
- To control sampling error, commercial farms can be stratified by type and size of livestock they maintain
- Those that manage mixed livestock may be put in a different stratum
- Very large or specific holdings can be put in a take-all stratum where all units will be surveyed
- Selection of farms may be done using SRS or systematic sampling

3.3 Sampling and estimation techniques

Stratification & Domains

- In a small holdings' survey, the stratification measures can be:
  - geographic characteristics (e.g. districts, ecological zones)
  - livestock system
  - economic characteristics e.g. number of small holdings that raise cattle, or those that raise poultry

- To determine which stratification measures are effective, choose stratification variables
  (i) Highly correlated with main survey variables (e.g. total number of cattle)
  (ii) With values available for all PSUs
  (iii) Stable and consistently measurable across all PSUs

3.3 Sampling and estimation technique

Stratification & Domains

- Survey estimates are needed at national & subgroups e.g. regions, provinces or socio-economic subdivisions (such as agro-ecological zones)
- The **sample** is designed to provide reliable estimate for important subgroups/divisions (referred as sampling domains)
- In most cases, geographic subdivisions sampling domains are used
- Sample size will be determined at the sampling domains level
- The total sample size across domains increases as number of domains increases

3.3 Sampling and estimation techniques

Developing Sampling Frame for Small Holdings

- Subsistence and small specialized livestock farmers comprise the majority of the population units in developing countries
- Census i.e. population & agriculture or livestock is best data source to identify the areas where farmers live
- Could be used to construct the sampling frame of small holdings or commercial holdings from which to select the sample for the surveys
- PSUs can be constructed with their corresponding measures of size

3.3 Sampling and estimation techniques

Developing Sampling Frame for Small Holdings

- PSUs could be enumeration areas from a previous census, a village or group of villages or otherwise defined settled area, depending on the appropriate cluster size

- Data from the administrative reporting system could be used to construct a sampling frame of PSUs in absence of census data

- Livestock holders can be listed in selected PSUs, from which sample livestock holders can be drawn

3.3 Sampling and estimation techniques

Developing Sampling Frame for Commercial Holdings

- Commercial livestock producers are listed in the census of agriculture and/or in the census of establishments, in case they are also businesses

- In some countries, commercial holdings can be operated by households

- Usually belong to a specific commodity business organization
  - listed in government regulatory bodies (e.g. Bureau of Animal Industry) for licensing and monitoring purposes

- Data sources can be consolidated to build an updated sampling frame of commercial livestock producers

Now lets look at . . .

Determining Sample Size

3.3 Sampling and estimation techniques

Method/Approach 1

Set the level of precision for the survey such that

$$Py\text{-}Y<e=1\text{-}\alpha$$

Where y is the sample mean of the characteristic of interest y, e is the margin of error, and 1-α is the confidence level

3.3 Sampling and estimation techniques

To obtain the level of precision *e*, the sample size under SRS, should be:

$$n_{SRS} = \frac{z_{\alpha/2}^2 S^2}{e^2 + \frac{z_{\alpha/2}^2 S^2}{N}}$$

where

$z_{\alpha/2}$: standard normal abscissa at α/2

S^2 : the population variance of the characteristic of interest y

N : the population size

N and S^2 are derived from the sampling frame (Lohr, 2010)

3.3 Sampling and estimation techniques

Method/Approach 1

If N is quite large, or if it is unknown, the sample size may be approximated to:

$$n_{SRS} \approx \frac{z_{\alpha/2}^2 S^2}{e^2}$$

3.3 Sampling and estimation techniques

Method/Approach 2

The maximum expected coefficient of variation of the sample mean y may be set such that

$$\mathrm{CV}(\bar{y}) \leq \delta$$

and hence,

$$n_{SRS} > \frac{[\mathrm{CV}(y)]^2}{\delta^2}$$

where

$\mathrm{CV}(y)$: the population CV of the characteristic of interest y and may be computed from the sampling frame

3.3 Sampling and estimation techniques

Method/Approach 2

However, for the latter, since the design involves clusters, the final sample size will be determined by multiplying the obtained sample size by the design effect

$$d^2(\bar{y}) = \frac{\mathrm{var}(\bar{y})_{complex}}{\mathrm{var}(\bar{y})_{SRS}}$$

3.3 Sampling and estimation techniques

Deciding the number of households at PSU level

For small holdings survey, the optimal number of livestock households need to be determined per PSU

$$b_{opt} = \sqrt{\frac{C_1}{C_2}\frac{(1-\hat{\rho})}{\hat{\rho}}}$$

with C_1 : cost of adding an additional PSU into the sample

C_2 : cost of an additional interview (Kish,1965)

3.3 Sampling and estimation techniques

Deciding the number of households at PSU level

- The intra-class correlation can be estimated from previous livestock surveys by

$$\hat{\rho} = \frac{d^2 - 1}{b - 1}$$

- with d^2: the design effect

  b: the number of smallholder households from previous survey

3.3 Sampling and estimation techniques

Deciding the number of households at PSU level

- In case that there is no previous survey, the intra-class correlation can be estimated from the sampling frame:

$$\rho = 1 - \frac{\bar{B}}{\bar{B}-1} \frac{\sum_{\alpha}\sum_{\beta}\left(y_{\alpha\beta} - \bar{y}_{\alpha}\right)^2}{\sum_{\alpha}\sum_{\beta}\left(y_{\alpha\beta} - \bar{y}\right)^2},$$

A = total number of PSUs for a given domain

$y_{\alpha\beta}$ = the characteristic of interest for livestock smallholder household β in PSU α

$$\bar{y}_{\alpha} = \sum_{\beta} y_{\alpha\beta} \Big/ B_{\alpha} \qquad \bar{y} = \sum_{\alpha}\sum_{\beta} y_{\alpha\beta} \Big/ \sum_{\alpha} B_{\alpha} \qquad \bar{B} = \sum_{\alpha} B_{\alpha} \Big/ A$$

3.3 Sampling and estimation techniques

Deciding the number of households at PSU level

- C_1 and C_2 should be determined based on the budget for the conduct of similar previous surveys

- C_1 covers:
  - travelling expenses
  - per diem including the allowance given for gasoline expenses of those with official motorcycles during enumeration and supervision.

- C_2 on the other hand, can be determined by dividing the daily wage of hired interviewers by the minimum daily output derived from previous surveys. The minimum daily output can also be derived from the pre-test information, if any.

3.3 Sampling and estimation techniques

Sample Selection

- **For commercial holdings**
  - Simple random sampling (SRS) after determining sample size and stratification
  - You can take all holdings (if few)
  - Use systematic sampling as in steps below

Steps

1) Sort all farms in a stratum according to an auxiliary variable that has not been used for stratification
2) Select farms using a random number drawn between 1 and the sampling interval which is the ratio between the total number of farms in the stratum and the number of farms that should be in the sample.
3) The first PSU in the sample is the first PSU that has a cumulative measure of size larger or equal to the sampling interval
4) Add the sampling interval to the random start to get the next farm in the sample and so on, until the total number of farms to be sampled is achieved

3.3 Sampling and estimation techniques

Sample Selection

- **For small holdings**
  - PSUs will be selected first using PPS
  - More preferable is systematic PPS over PPS to ensure a more balanced set of PSUs

In systematic PPS, PSUs can be sorted according to relevant auxiliary variable(s) before sample selection to ensure that the PSUs selected will be representative of the population distribution of PSUs

3.3 Sampling and estimation techniques

Sample Selection

- **For small holdings (continued)**
  - To select the PSUs in systematic PPS in a stratum, the sampling interval will be the ratio of the total measure of size in the stratum over the number of PSUs to be selected
  - Draw a random number between 1 and the computed sampling interval is then drawn
  - The first PSU in the sample is the first PSU that has a cumulative measure of size larger or equal to the sampling interval
  - The next PSU for the sample is that with the cumulative measure of size that is larger than the sampling interval plus the random number drawn, and so on
  - Small holdings will then be selected in the sampled PSUs

3.3 Sampling and estimation techniques

Survey weights and estimation

- **<u>Steps</u>**
  1) Determine each unit's initial weight i.e. inverse of unit's selection probability
  2) Adjust the initial weight to consider possible nonresponse
  3) Adjust the initial weight to make the estimates conform to some known population totals
  4) Compute final survey weight to each responding unit as a product of the initial weight, the nonresponse adjustment, and the population weighting adjustment

Use final survey weights in all analyses to produce valid estimates of population parameters

3.3 Sampling and estimation techniques

Survey weights and estimation

- For small holdings raising livestock, the most common form of nonresponse weighting adjustment is a weighting class adjustment
- The full sample of respondent farmers and non-respondent farmers is divided into a number of weighting classes or cells and nonresponse adjustment factors are computed for each cell *c* as

$$w'_c = \frac{\sum_{i \in rc} w_{di} + \sum_{j \in mc} w_{dj}}{\sum_{i \in rc} w_{di}} = \frac{\sum_{i \in sc} w_{di}}{\sum_{i \in rc} w_{di}}$$

- The denominator of W'_c is the sum of the weights of respondent farmers (indexed r) in *cell c.* The numerator adds together the sum of the weights for respondent farmers and the sum of the weights for eligible non-respondent farmers (indexed m for missing) in cell *c*

3.3 Sampling and estimation techniques

Survey weights and estimation

- The two sums in the numerator give the sum of the weights for the total eligible sample (indexed s) in cell *c*
- Thus the nonresponse weight adjustment **w'c** is the inverse of the weighted response rate in *cell c*
- A low response rate results in a large non response adjustment for the cell,
- This significantly lowers the precision of survey estimates
- Attempts must be made to avoid high non response

3.3 Sampling and estimation techniques

Estimation of the population total Y

- The total is estimated

$$\hat{Y} = \sum w_i y_i$$

- $\hat{Y}$ is the estimate of population total, $\boldsymbol{y}_i$ the value of variable y and $\boldsymbol{w}_i$ the final weight for respondent farmer ***i***
- It is possible to do population total estimate by subgroups using:

$\boldsymbol{y}_i$ = 1 if farmer ***i*** is in the desired subgroup

0 otherwise

3.3 Sampling and estimation techniques

Estimation of the population mean

- The population mean is estimated by:

$$\bar{y} = \sum w_i y_i \,/\, \sum w_i$$

- $\bar{y}$ is the estimate of population mean, $\boldsymbol{y_i}$ the value of variable y and $\boldsymbol{w_i}$ the final weight for respondent farmer ***i***

3.3 Sampling and estimation techniques

Estimation of the population ratio

- Ratios is estimated by $R = Y/X$

where X = the population total for another variable denoted by x

- The ratio R may be estimated by the ratio:

$$r = \sum w_i y_i \,/\, \sum w_i x_i$$

Conclusion

- In this session, we have discussed:
  - Data sources
  - Data collection methods
  - Survey design and survey implementation
- Country context matters in collecting livestock production and productivity indicators
  - ➢ Depend on country specific needs, type of agricultural holdings, respondent ability etc.

Reference

Upcoming Publication (2018)

Global Strategy to improve Agricultural and rural Statistics (2018). *Guidelines on methods for Estimating Livestock Production and Productivity*

Available online : http://gsars.org/

Thank You