The Principal steps in a Sample Survey :

Conducting a large scale sample survey involves two main stages —

1) Planning and execution of sample survey
2) Analysis of data and reporting

1) Planning and Execution :

The following are some of the main steps involve in the planning and execution of sample surveys —

a) Objectives → A lucid statement of the objectives is most helpful, without this it is easy in a complex survey to forget the objectives when engrossed in the details of planning and to make decisions.

b) Population to be Sampled → The word population is used to denote the aggregate from which the sample is choosen. In sampling a population of farms, rules must be set up, to define a farm; and these rules must be usable in practice. The population to be sampled should coincide with the target population [The popⁿ about which information is wanted]. Sometimes for reasons of practicability the sampled population

is more restricted than the target population.

c) Data to be collected → Data relevent to the purpose of the survey should be collected only. If there are too true many questions, the respondants begin to lose interest in answering them. On the other hand it must be ensured that no important item is missing. A practical procedure is to prepare outlines of the tables that the survey should produce.

d) Degree of precision desired → The results of sample surveys are always subject to some uncertainity because, only a part of the population has been measured and the measurements are always subject to some error. The uncertainity can be reduced by taking larger samples and by using superior instruments of measurement. Consequently, the specification of the degree of precission wanted in the results is an important step.

e) Method of collecting information : The method of collecting the information [wheather by mail or by interview or otherwise] has to be decided, keeping in

view the cost involved and the accuracy aimed at. Mail surveyes cost less but there may be considerable non response. Interviewers cost more and there are interviewer errors, but without the interviewers the data collected may be worthless.

f) The frame and the Sampling Unit(s) → Before selecting the sample the population must be divided into parts that are called sampling units. These units must cover the whole population and they should not overlap, in the sence that every element in the population belongs to one and only one unit. For, e.g, in sampling the people the unit might be an individual person, the members of a family etc. The construction of this list of sampling units called a frame, is often one of the major practical problems. In order to cover the population to be sampled, there should be some list, map or other acceptable material. [called the frame], which survves as a guide to the population to be covered.

g) Selection of the Sample: At this stage the question of the size of the sample, the manner of selecting the sample and the estimation of the population characteristics along with their margin →

of uncertainity are some of the technical problem that should recieve the most careful attention.

b) Questionere or shedule ⟶ The questionere [to be filled in by the respondant] or shedule [to be filled by interviewer] forms a very important part of the sample survey. Having decided on the data to be collected, the problem of their presentation requires considerable scheme.

A schedule contains a list of items on which information is sought but the exact form of the questions to be asked is not standerdized but left to the judgement of the enumerators. A questionnai on the other hand is a set of questions that could actually be put to the informants in a specified order. While either of this may be used in an interview type of enquiry [the general distinction being that a shedule will be fielding by the enumerator while a questionaire will be filled in by the informant himself], a mail questionnaire type of enquiry necessarily uses the later..

i) Training of interviewers and their supervision :→

The success of a survey using the interview method depends largely on the ability of the interviewers to collect acceptable responces. Their selection and training is very important. Observation by a supervisor during the course of an actuall interview is valuable for maintaining standards.

2) Analysis of Data and Reporting ⟹

This consists of the following steps —

a) Scruting of Data → The first step is to edit the completed quiestionaire, in the hope of amending recording errors, or atleast of deleting data that are obviously ~~excorrect~~ erronious.

b) Tabulation of data → Wheather hand tabulation or mechanical tabulation is to be taken recourse to depends upon the quantity of data. For a large scale survey involving several thousands of individuals, machine tabulation is expected to be more economical and quicker.

c) Statistical Analysis → The tables may be further utilised for deriving necessary estimates for population

→

characteristics or for testing hypothesis, if any. Different methods of estimation may be available for the same data.

d) Reporting → The report should incorporate a detailed statement regarding all the stages of the survey and should present all the statistical information collected. The data should be properly interpreted the necessary conclusions should be derived and the right recomendation should be made. The technical aspects of the design of the survey, for example, the types of estimators used and their margins of errors must be mentioned.

e) Information gain for future surveyes → Any completed sample is potentially a guide to improve future sampling, in the data that it supplies about the means, standar deviations, and nature of the variability of the principle measurements and about the costs involved in getting the data.

17/6/14

* Sampling and Non-Sampling Errors :

The errors involve in the collection, processing and analysis of data may be broadly classified under the following two catagories —

1) Sampling errors
2) Non-sampling errors

1) Sampling Errors :

Sampling errors have their origin in sampling and they arise due to the fact that, only a part of the population, i.e, sample has been used, to estimate population parameters and to draw inferences about the population. As such, the sampling errors are absent in complete enumeration. Sampling errors are primarily there due to the following reasons —

i) Faulty selection of the sample → some of the bias is introduced by the use of defective sampling techniques for the selection of a sample. E.g : purposive or judgement sampling in which the investigator deliberately selects a representative sample to obtain certain results. This bias can be overcome by strictly adhering to a simple random sample or by selecting →

17/6/14

* Sampling and Non-sampling Errors :

The errors involve in the collection, processing and analysis of data may be broadly classified under the following two catagories —

1) Sampling errors
2) Non-sampling errors

1) Sampling Errors :

Sampling errors have their origin in sampling and they arise due to the fact that, only a part of the population, i.e, sample has been used, to estimate population parameters. and to draw inferences about the population. As such, the sampling errors are absent in complete enumeration. Sampling errors are primarily there due to the following reasons —

i) Faulty selection of the sample → some of the bias is introduced by the use of defective sampling techniques for the selection of a sample. E.g : purposive or judgement sampling in which the investigator deliberately selects a representative sample to obtain certain results. This bias can be overcome by strictly adhering to a simple random sample or by selecting →

a sample at random subject to restrictions. Which, while improving the accuracy are of such nature that they do not introduce bias in the results.

ii) Substitution → If difficulties arise in enumerating a particular sampling unit includ in the random sample, the investigators usually substitute a convenient member of the population This obviously leads to some bias since the characteristics possessed by the substituted unit will usually be different from those by the unit originally included in the sample.

iii) Faulty Demarkation of sampling units → Bias due to defective demarkation of sampling units is particularly significant in area surveys such as, agricultural experiments in the fields or crop cutting survey. In such surveys while dealing with border line cases, it depends more or less on the discretion of the investigator, whether to include them in the sample or not.

iv) Constant error due to the improper choice of the statistics for estimating population parameters → If $x_1, x_2, \ldots, x_n$ is a sample

of n independent observations then the sample variance $s^2 = \left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]$ is biased. As an estimate of the population variance whereas the statistic $\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]$ is an unbiased estimator of σ^2.

Note: Increase in the sample size usually results in the decrease in sampling error. Infact this decreasing sampling error is inversly proportional to the square root of the sample size.

2) Non-Sampling Errors:

As distinct from sampling error, ~~non-sampling errors~~ which are due to the inductive process of infering about the population on the basis of a sample, the non-sampling errors primarily arise at the stage of observation ascertainment, and processing of the data and thus present in both complete enumeration and the sample survey. Thus, the data obtained in a complete census, although free form sampling errors, would still be subject to non-sampling errors; whereas, data obtained in a sample survey would be subject to both sampling and non-sampling errors.

Non-sampling errors can occur at every stage of the planning or execution of census or sample survey. The preparation of an exosting list of all the sources of non-sampling erro is a very difficult task. However a careful examination of the major ~~fases~~ phases of survey (complete or sample survey), indicates that some of the important non sampling er arise from the following factors. ——

i) Falty planning or definitions → The pla of a sample survey consists in explecitely stating the objectives of the survey. These objectives are then translated into a set of characteristics for which data are to be collected and a set of ~~spef~~ specifications for collection, processing and publishing. Here the non sampling errors can be due to ——

a) Data specification being inadequate and inconsistent with respect to the objectives of the survey.

b) Error due to the location of the unit and actual measurement of the characteri errors in recording the measurements,

errors due to ill-designed questionnaires etc.

c) lack of trained and qualified investigators and lack of adequate supervisory staff.

** ii) Response Errors → These errors are introduced as a result of the responses furnished by the respondents and may be due to any of the following reasons —

a) Response errors may be accidental, E.g., the respondents may misunderstand a particular question and accordingly furnish improper information unintentionally.

b) Prestige bias: An appeal to the pride or prestige of persons interview may introduce another kind of bias, called prestige bias. By virtue of which he or she may upgrade his education, IQ, income etc. or downgrade his/her age. Thus resulting in wrong results.

c) Self Interest: Quite often in order to safe guard one's self interest one may give incorrect information. E.g.; a person may be give an underestimate of his saving or production and an over statement of his expenditure or requirements etc.

d) Bias due to interviewer: Sometimes the interviewer by may affect the accuracy of the →

response by the way he asks questions or records answer. The information obtained on suggestion from the interviewer is likely to be influenced by interviewer's believes and prejudices.

e) Failures of Respondant's memory : One source of error which is common to most of the methods of collecting informations is that of 'recall' many of the questions in surveyes refer to happenings or conditions in the past and there is a problem, both of remembering the event and ~~enumerating~~ associating it with the correct time period.

**

iii) Non - Response bias → Non - response biases occur if the full information is not obtained for all the sampling units. In house to house survey non response usually results if the respondant is not found at home even after repeated calls or if he is unable to ~~farnis~~ furnish the information on all the questions or if he refuses to answer certain questions. Therefore, some bias is introduced as a consequence of the exclution of a section of the population with certain peculiar char

due to non response.

** iv) Errors in Coverage → If the objectives of the survey are not precisely stated in clear cut terms then it may result in the inclusion in the survey of certain units which are not to be included or the exclusion of certain unit which have to be included in the survey under the objectives. Eg: in a census to determine the number of individuals in the age group 20 to 50 years, more or less serious errors may occur in deciding whom to enumerate unless particular community or area is not specified and also the time at which age is to be specified.

v) Compiling Errors → Various operations of data processing such as editing and coading of the responses, ~~punchi~~ punching of cards, tabulation and summerising the original obsⁿ made in the survey are a potential source of errors. Compilataion errors are subject to controll through verification, consistency check etc.

vi) Publication errors → Publication errors, i.e, the errors committed during presentation and

printing of tabulated results are basically d
to two sources. The first reffers to the mech
of publication — the proofing error and th
like. The other which is of more seri
nature lies in the failure of the survey
organisation to point out the limitations
the statistics.

21/7/1

Remarks :

① In a sample survey non-sampling errors m
also arise due to defective frame and faulty
selection of sampling ~~units~~ limits.

② It is obvious that the non-sampling error
are likely to be more serious in a complete
census. As compare to sample survey, since in
sample survey the non sampling errors can be
reduced to be a better extend by qualified, tr
and experienced personnel better supervision an
better equipment for processing and analising
relatively smaller data as compared to a com
census.

It has already being pointed out that usually
sampling error decreases with the increase in
sample size.

Only other hand as the sample size increases
non-sampling errors tends to increase.

(3) Quite often the non-sampling error in a complete census is greater than both the sampling and non-sampling error taken together in a sample survey. Obviously in such situation sample survey is to be preferred to complete census.

Basic Principle of Sample Survey :

The two basic principle for the design of a sample survey are —

a) Validity
b) Optimisation

It takes into account the factors of —

i) Precision or efficiency
ii) Cost

a) Validity :

By validity of a sampling design we mean that the sample should be so selected that the results could be interpreted objectively with certain confidence or in terms of probability.

In other words validity of a sampling design ensure that valid estimates or states about the population characteristic should be available for this. It is necessary to attach ~~member~~ probability to each member of the population to be included in the sample.

→