A Practical Guide to Sample Surveys

香港統計學會
Hong Kong Statistical Society

Y.K. Chan
K.W. Ng
F.W.H. Ho
S.M. Shen

A PRACTICAL GUIDE

TO

SAMPLE SURVEYS

~~~~~~

Y. K. Chan
Department of Sociology
The Chinese University of Hong Kong

F. W. H. Ho
Census and Statistics Department
Hong Kong

K. W. Ng
Department of Statistics
University of Hong Kong

S. M. Shen
Department of Statistics
University of Hong Kong

香港統計學會
Hong Kong Statistical Society
~~~~~~

Correspondence address:

Hong Kong Statistical Society,
c/o Department of Statistics,
University of Hong Kong,
Pokfulam Road,
Hong Kong.

FOREWORD

Sample surveys have become very popular these days in Hong Kong and are used for a wide range of purposes. It is expected that this surging trend for the widespread use of sample surveys will continue in the years to come. However, it should be emphasized that those surveys which are not properly conducted do not produce meaningful findings. For surveys conducted privately for business purposes, the results may lead to unsound business decisions but still they affect only individual firms. For surveys whose results are to be publicised and may have implications on public policies, it is a social responsibility of the sponsors of such surveys to ensure that they are properly conducted.

The Hong Kong Statistical Society has been actively advocating the proper conduct of sample surveys through various activities. The Society published a brochure entitled 'What Is A Survey?' (first printing in 1982 and second printing in 1984) and a press release in 1987 listing the important points to note in conducting surveys. Seminars on sample surveys were organized by the Society on various occasions. A course has also been instituted in conjunction with the Department of Extra-mural Studies, University of Hong Kong.

The Editorial Board of the Hong Kong Statistical Society feels that it is desirable to have a more detailed booklet giving advice on the proper conduct of sample surveys. We have thus invited members of the Society, Dr. Y.K. Chan of The Chinese University of Hong Kong, Mr. F.W.H. Ho of the Census and Statistics Department, Hong Kong, Dr. K.W. Ng and Dr. S.M. Shen of the University of Hong

Kong to prepare this booklet 'A Practical Guide to Sample Surveys'. We appreciate the authors for their contribution and it is hoped that the booklet would be of assistance to survey-takers and users of survey results.

The Editorial Board is also grateful to the ex-Secretary, Mr. C.M. Wong, Department of Community Medicine, University of Hong Kong, for his service during the preparation of this booklet.

Editorial Board,
Hong Kong Statistical Society

March 1991.

ACKNOWLEDGEMENT

We would like to thank the Editorial Board of the Statistical Society of Hong Kong for giving us the opportunity to meet the challenge of preparing a booklet which deals with a complicated subject in a simple manner. During the preparation, we have received many valuable comments on earlier drafts from the Editorial Board and members of the statistical profession in Hong Kong. We are especially grateful to Mr. George Ng, Census and Statistics Department, Hong Kong, Dr. H.W. Mui, City Polytechnic of Hong Kong, Dr. Stephen L.C. Chan, Hong Kong Baptist College, and Dr. Joseph M. Chan, The Chinese University of Hong Kong.

We are happy to acknowledge the skilful and efficient word processing service of Miss Irene Cheung, Department of Statistics, University of Hong Kong.

Y.K. Chan
F.W.H. Ho
K.W. Ng
S.M. Shen

March 1991.

A PRACTICAL GUIDE TO SAMPLE SURVEYS

CONTENTS

CHAPTER 1
OVERALL PLANNING OF A SURVEY OPERATION

1.1 Defining the objective

The word *survey* is most often used to describe a method of gathering information from a number of individuals (i.e. a sample) in order to learn something about the population from which the sample is to be drawn. The first step in planning a survey is to set out the objectives of the investigation. This is generally the task of the sponsor of the inquiry.

It is very important that the objectives of the survey are clearly defined. It should be as specific, clear-cut and unambiguous as possible.

1.2 Quality required of the information to be obtained

The survey-taker should ensure that the information obtained in the survey is good enough to meet the pre-determined objectives.

To be useful, the statistics need not be fully exact, but they do need to be sufficiently reliable to serve the particular needs. No overall criterion of reliability applies to all surveys since the margin of error that can be tolerated in a study depends on the analysis and precision required for meaningful actions or recommendations that will be influenced by the data collected. There are indeed occasions when a relatively high margin of error is acceptable. For example, in a survey to study the general housing situation in terms of whether there is a tight housing supply, if the true vacancy rate is very low, say one percent, survey results that show double this percentage (100% error) will not do any significant harm; and results in the range of zero to three percent will lead to the same conclusion — a tight housing market. However, if the purpose is to actually carry out planning on the number of units to be built, then we would require a more accurate assessment of the situation.

No general rule can be laid down to determine the reliability that would apply to all surveys. It is necessary to consider the purpose of the particular study, how the data will be used, and the effect of errors of various sizes on the action taken based on the survey results. These factors will affect the sample size, the design of the questionnaire, the effort put into training and supervising the field staff, and so on. Estimates of error should also be considered in analyzing and interpreting the results of the survey. Usually, it can be expected that the cost of greater accuracy may require more resources requirements and longer time.

1.3 Defining the population

The term *population* used here means the totality of elements (units of analysis) under study. For a given survey objective, the elements may be persons; but for other objectives, they can alternatively be households, business firms, schools or any other units.

The population covered by the survey should be precisely and carefully specified and clearly defined according to the survey objectives, because the interpretation of survey results will depend on how it is defined. Consider, for example, a survey to be conducted in Tuen Mun to discover the degree of support of local inhabitants for the introduction of a new Tuen Mun – Yuen Long bus service. In defining the population of the survey, there are a number of relevant questions that we should ask: What is the minimum age for the persons to be surveyed? Should persons living temporarily in Tuen Mun be excluded, and in fact, how are such persons defined? In practice, a variety of questions like these arise in defining most populations, making the definitional task less straightforward than it might at first appear.

In general, it is useful to firstly define the ideal population required to meet the survey objectives — the *target population.* This ideal definition can then be modified to the *survey population* to take account of any practical constraints. The major advantage of starting with the ideal target population is that any exclusions and constraints suffered in formulating the subsequent survey population are explicitly identified, thus enabling the magnitude and consequences of the restrictions to be assessed.

1.4 Conditions for obtaining proper response

A survey can produce meaningful results only in so far as the respondents are both willing and able to provide the relevant information. In this regard, actions should be taken to boost respondent cooperation.

Besides, the survey-taker must pay attention to whether the contents of the survey are too sensitive, whether they unduly invade the respondent's privacy, and whether they are too difficult even for a willing respondent.

1.5 Professional input and general resource requirements

In view of the technical nature of surveys, professional input is required in the conduct of a survey. This is required right from the start, since rectification is normally very difficult to make when the survey has been completed and its results are judged to be invalid.

To ensure that a well-conceived survey design can be implemented smoothly, adequate resources such as time, money and manpower must be made available for the survey. The fact that information is usually a costly commodity should not be overlooked. With such cost implications, whether to seek information or not hinges on the usefulness of the information.

References for this chapter can be found in Moser and Kalton, Chapters 1, 2 and 3, and Babbie, Chapters 3 and 4.

CHAPTER 2
DESIGN AND SELECTION OF SAMPLES

2.1 Use a probability sample

Sample design is both an art and a science concerning how to select the part of the population to be included in a survey. A basic distinction is whether the selected sample is a probability sample or not.

With a *probability sample*, each element has a known, nonzero chance of being included in the sample. Consequently, subjective selection biases are avoided, and statistical theory can be used to derive properties of the *survey estimators* (the mathematical formulas for computing the statistics we want from the sample observations).

Non-probability sampling covers a variety of procedures, including the inclusion of respondents who volunteer to respond without being requested, haphazard sampling (choosing whoever is convenient) and judgemental sampling (subjective choice of elements for the sample on the grounds that they are considered 'representative' of the population). The weakness of all non-probability sampling is the absence of a theoretical framework for statistical inference. Validity and margin of error of survey results can be assessed only by subjective evaluation, not by sound statistical methods.

Hence, as far as practicable, probability sampling should be used rather than non-probability sampling. This makes it scientifically valid to draw inferences from the sample results about the entire population which the sample represents.

It is necessary to note that some sampling may appear objective, although not really so. One obvious example is *quota sampling*, in which the interviewers are instructed to find and enumerate specified quotas of 'representative' sample elements roughly proportional to the population on a few control variables such as age, sex or geographic areas.

For example, the latest population distribution of Hong Kong by broad geographic area by sex is available from the 1986 population By-Census as shown in the following table. Based on this information, the quotas in a quota sample of 1,000 adults to be included in a household opinion survey can then be computed so that 109 males and 109 females should be selected from the Hong Kong Island, 179 males and 170 females should be selected from the New Territories and so on.

Broad Area	Sex	No. of persons	(%)
Hong Kong Island	M	859,267	10.9
	F	856,593	10.9
Kowloon and New Kowloon	M	1,195,663	22.2
	F	1,106,028	20.5
New Territories	M	966,245	17.9
	F	914,921	17.0
Marine	M	21,289	0.4
	M	15,991	0.4

The sample is then divided among the interviewers, who are supposed to do their best to find and enumerate household adults who fit the restrictions of their quota controls in the survey (e.g. 179 male household adults in New Territories). By assigning quotas it was hoped to avoid, or at least to control, the subjective selection biases that would occur if the interviewers were given a free hand in their choice of respondents. But quota sampling still falls outside the domain of probability sampling because, by leaving the selection of units to interviewers, it does not bring about randomization within each class. Selection within the quotas is thus haphazard or subjective.

2.2 Consider sampling error

Apart from other sources of survey error, statistics derived from samples are subject to sampling error. Sampling error arises because the particular estimate obtained in the survey is only one among the many possible estimates that could have been obtained by using the same sample design and sample size.

For example, suppose we want to estimate the proportion (P) of the population which have a certain characteristic by using the sample proportion (p) obtained in the survey.

In this connection, there is one important question that we can ask: Given a sample design and sample size, what values of p are possible, and what is the probability of occurrence of each of these possible p's? The array of all possible values of p, each with its probability of occurrence, is called the *sampling distribution* of the possible p's for a fixed population, sample design, and sample size. This sampling distribution represents the random fluctuation of p due to the specific sample design used by the survey-taker. The variability of p over all possible samples of the same design and size is measured by its *standard error*, which is the standard deviation (a measure of dispersion of statistical distribution) of the sampling distribution of p.

2.3 Determination of sample size

The required size of the sample depends on the choice of the method of sampling and the desired level of precision (or, equivalently, the acceptable *margin of error* or *sampling error*) of the estimates. It should be worked out using established statistical methodology. The following example will illustrate how the sample size may be determined in a practical context.

Suppose a face-to-face interview survey is to be conducted to estimate the percentage of the population of a certain district (with population of say about 15,000 adults) who say they would make use of a new library if one were built. To determine an appropriate sample size, it is first necessary to specify the degree of precision required for the estimate. This is no easy task, and initially the degree of precision required is often overstated. Suppose, for instance, the initial specification calls for an estimate that is within 2 percentage points of the true population proportion with 95% probability. In other words, the 95% confidence limits should be the sample proportion plus or minus 0.02. This specification thus requires that

$$1.96 \times \text{s.e.}(p) = 0.02 \qquad (2.1)$$

where p is the sample proportion and s.e.(p) is the *standard error* of p (see §5.2).

Assuming initially the use of simple random sampling (SRS), and ignoring the *finite population correction* (fpc) term for the time being, we learn from standard statistical theory that

$$[\text{s.e.}(p)]^2 = PQ/n' \qquad (2.2)$$

where P is the true population proportion, Q is equal to $(1 - P)$, n' is the initial estimate of the sample size. From mathematics, we know that PQ is largest at $P = Q = 0.5$, so a conservative guess is to set P equal to a value as close to 0.5 as is likely to occur. Let P be 0.4 then from equations (2.1) and (2.2), we have

$$n' = 2305$$

In the event one has no idea at all about P, set it to 0.5 for the calculation of n'.

If the initial sample size were small compared with the population size N, so that the fpc term could be ignored, it would be the required

sample size n. But since $n' = 2305$ is not small relative to the population size of the district in question ($N = 15,000$), we must take the fpc term into account. From statistical theory, we can obtain n by the following formula

$$n = Nn'/(N + n'). \qquad (2.3)$$

In this case $n = 1998$.

Another factor needs to be included in the calculation of the required sample size is potential non-response. Suppose that the response rate is expected to be 75% (say from previous experience). Then the selected sample of 1998 adults has to be set at $\frac{1998}{0.75} = 2664$. This adjustment does not serve to address the problem of 'non-response bias' that we are going to discuss later, but serves to produce the required sample size to meet the desired precision of p.

Having reached this point, the survey-taker may decide to review the initial specification of precision to see if it can be relaxed. Since the level of precision required for an estimate is seldom fixed, the sample size is usually determined from a rough-and-ready assessment of survey costs relative to the level of precision that will result.

It should be noted that the selected sample size depends on predictions of a number of quantities, such as the proportion of the population who say they would use the library (P), and the non-response rate. Errors in predicting these quantities cause the estimate to have a level of precision different from that specified, but the estimate remains a reasonable estimate of the population parameter.

The determination of the sample size for more elaborated methods of sampling will depend on the standard error formulas of the estimates, such as those given in §5.3.

2.4 Some basic probability sample designs

Simple random sampling, which may take slightly different forms in practice, is basically sampling by lot-drawing.

Another commonly used random sampling method is *Systematic Sampling* which is the familiar '1 in k' method, i.e. pick a random start (say the mth unit in an ordered queue of the members of the population, with $m \leq k$), and take the $(m + k)$th, $(m + 2k)$th, $\cdots$ units into the sample.

In many situations a certain amount of supplementary information is known about the elements of the population to be studied. This supplementary information can be used to improve the sample design through the technique of *stratification*. In essence, stratification is the classification of the population into non-overlapping subpopulations, or strata, based on some supplementary information. The classification is done in a way such that units belonging to the same stratum are relatively more homogeneous with respect to a certain characteristic(s) which is highly related to the information to be collected in the survey. Independent samples are then selected from each of the strata. The benefits of stratification arise from the fact that the sample sizes in the strata are controlled by the sampler, rather than being randomly determined by the sampling process. It enables the survey-taker to obtain more precise estimates (as statistical theories show), and to employ different sampling and data collection methods within different strata.

In most sampling situations the population can be usefully regarded as being composed of a set of groups of elements. As discussed above, one sampling use for such groups is to treat them as strata, in which case separate samples are selected from each and every group. Another sampling use is to treat them as clusters, in which case only

a sample of such clusters is included in the survey. (To enable sampling errors to be assessed at least two clusters should be selected.) If all the elements in selected clusters are included in the sample, the method is called *cluster sampling*. If only a sample of elements is taken from each selected cluster, the method is called *Two-stage sampling*. In more complex sampling situations, often a hierarchy of clusters is used. For example, in a survey of students in Hong Kong we may first select a sample of schools, then a sample of classes within each selected school, and finally a sample of students within each selected class. This general method is called *multi-stage sampling*.

In general, simple random sampling serves as a useful benchmark against which to compare other more complex sample designs. Suppose that in order to achieve a given desired level of precision in the final estimates, a simple random sample of size equal to n is required. If for some reasons the survey-taker uses a cluster sample design, then usually a sample size larger than n would be needed. On the other hand, if he uses an appropriate stratified sample, then it is likely that sample smaller than a n would suffice. Among other factors, it should be noted that the choice of sampling method will have an important implication on the sample size required to achieve a given desired level of precision in the final estimates.

2.5 Using a good sampling frame

An essential requirement for any form of probability sampling is the existence of a *sampling frame* from which the sample members can be selected. The survey-taker should therefore try his best to secure a good sampling frame of the population from which the sample is selected.

When a list of all the population elements is available, the frame may be just the list. When there is no list, the frame is some equivalent procedure for identifying the population elements. For example, in area sampling, each element of the population may be associated with a particular geographical area called area segment, with well-defined natural or artificial boundaries. People or households are associated with the area of their residence, or main residence if they have more than one, so that there is a one-to-one matching between people and area segments. Then a sample of area segments is drawn from the total area, and an interviewer canvasses the sample area segments and lists all the appropriate units (e.g. households) so that some or all of them can be designated for inclusion in the final sample.

The sampling frame is a major ingredient of the overall sample design. It provides a means of identifying and locating the population elements, and it usually contains a good deal of additional information that can be used for stratification or clustering.

The ideal sampling frame would contain listing of each and every population element, once and once only, and would contain no other listings. In practice this ideal is seldom realized, and the survey-taker has to be on the lookout for imperfections. There are four types of potential frame problems: missing elements, clusters, foreign elements (or equivalently, blanks), and duplicate listings. These frame problems are discussed below.

(a) ***Missing elements*** :

Some population elements are not included on the frame. Missing elements may occur because a frame is either inadequate (that it is known not to cover the whole of the target population) or incomplete (that it actually fails to include some elements from the target population that it is supposed to cover). The distinction between inadequacy and incompleteness is of practical importance because the former category is often more easily recognized. For example, in a survey of

students in a school, the school list is inadequate if it has excluded part-time students although they are part of the target population for the study; the school list is incomplete if it is out-of-date and hence fails to include some new students.

Missing elements present the most serious frame problem because, unless a remedy is found, these elements have no chance of being selected for the sample, which thus fails to represent the total target population. Sometimes the problem may be sidestepped by re-defining the survey population to exclude the missing elements. This imperfect solution is often used when the excluded group is a negligible proportion of the total population, when the exclusion will have only minimal effect on the survey objectives, and when no simple alternative solution is available.

A preferred solution is to find supplementary frames to cover the missing elements, for example, special lists of part-time students and new entrants in the student survey in the above example. However, this solution may create the problem of duplicates because some students may appear on more than one list, but this lesser problem may be easily handled by the methods discussed under the frame problem of duplicate listings.

Often no suitable supplementary frame is available for the missing elements, and a solution involving some form of linking procedures may be sought. Linking procedures aim to attach missing elements to specified listings in a clearly defined way. When a listing is selected, any missing element linked to it is selected as well and the whole set of elements thus obtained is treated as a cluster. Linking thus gives rise to the frame problem of clusters, which may be handled by one of the methods discussed under the frame problem of clusters.

For example, in the student survey, the sampling frame comprises alphabetical lists of the students present at the original enrolment for each of the classes. A possible linking for missing students would then be to define each listing as representing the named student together with any student missing from the class list coming after that student and before the next listed student in the alphabetical order. To cover missing students at the start of the alphabet, the list may be treated as circular; thus, any missing student coming after the last listed student or before the first listed student is linked to the last student on the list. This form of linking is an example of what is known as a half-open interval. Another well known application is for sampling living quarters (LQs) from lists of LQs in street order, with each side of the street being taken separately. Using the half-open interval, missing LQs may be linked to the last listed LQ preceding them.

(b) Clusters :

Some listings refer to groups of elements, not to individual elements. As discussed above, the use of a linking solution for missing elements may create clusters. Clusters also occur in other circumstances, for example, when a sample of persons or households is required but the sampling frame is a list of living quarters.

One solution is to include all the elements in the selected clusters in the sample. This give the elements the same chance of appearing in the sample as their listings. If listings are sampled with equal probabilities, the elements are also sampled with equal probabilities. However, this take-all solution may lead to higher sampling error relative to a given sample size, or cause workload problems because the sample size becomes much larger than originally intended.

(c) Foreign elements (and blanks) :

Foreign elements are listings for elements which are outside the

scope of the survey, such as unemployed people in a survey of wage earners. Blanks are listings for elements that no longer exist in the population, such as living quarters that have been demolished but have not been removed for the list of addresses.

The method of handling blanks and foreign elements is straightforward: simply ignore them. The selection probabilities are not distributed in any way. As a result, the sample size may be smaller than the number of selections, since some blanks/foreign elements may be drawn and omitted. Hence, this should be taken into account in determining the sampling fraction, the ratio of sample size to population size, needed to generate the desired sample size.

A common error is to substitute the next element on the list for a selected blank/foreign element. This should not be done since it increases the selection probability for the next element (that next element could be selected either if it is selected directly or if the preceding blank is selected).

(d) Duplicate listings :

Some population elements have more than one listing on the frame. It often arises when the sampling frame is composed of several lists, and some elements appear on more than one list. The problem created by duplicates is that an element's selection probability varies with its number of listings.

One solution is to remove the duplicates from the whole frame, but this is often too expensive. Another solution is to employ the method of *unique identification*, associating each element with one of its listings in a clearly defined way (e.g. the first listing, or the oldest listing), and treating the other listings for that element as blanks. If the organization of the sampling frame, or the information it contains, does not readily permit the use of unique identification, then unique identification could still be applied during fieldwork. However, since a substantial proportion of the survey costs is used in making contact with respondents, it is uneconomical to drop some selections as blanks after interview. An alternative is to accept all selections and to use *weighting* in the analysis to adjust for the unequal selection probabilities. In this case, it is necessary to find out the number of duplicated listing of a sampled member during fieldwork (e.g. by asking the survey respondent certain appropriate questions).

2.6 Using proper weighting for unequal selection probabilities

In a good survey, it is not essential to use a sample design where the probabilities of selection of all individual respondents are equal. However, the probabilities of selection of different respondents must be known. In case of unequal probabilities of selection, it is necessary to ensure that proper weighting methods are applied to survey results to give unbiased population estimates.

Weights are used to assign greater relative important to some sampled elements than to others in the survey analysis. Generally, sample members with greater probability of selection are assigned proportionally smaller weights.

References for this chapter can be found in Moser and Kalton, Chapters 4, 5, 6 and 7, and Babbie, Chapters 5 and 6.

CHAPTER 3
DESIGN OF QUESTIONNAIRES

3.1 Good questionnaire is essential for quality data

In surveys, data are very often collected by interviews or mailed questionnaires, though sometimes by observations (Moser and Kalton, Chapters 10, 11 and 12). *Questionnaire*, a form or an instrument containing a set of organized questions for survey purposes, is employed to record information on survey respondents.

As we collect data by asking questions, we assume that answers given by individual respondents are correct. However, the correctness of answers depend on the ability and willingness of the respondents to give the right answers, respondents' understanding of the question wording, their accessibility to the required information, and the environment etc. all affect their ability to answer. People are usually reluctant to answer sensitive questions, embarrassing questions, or questions about their unpleasant experience, particularly in unsuitable interviewing setting or when they have doubt on the confidentiality and anonymity of the interview.

A well designed questionnaire is essential in guaranteeing data quality of surveys. When setting *opinion questions* in particular, be aware that answers to such questions are often sensitive to wording, emphasis, structure, and being subjective questions, which unlike objective questions, have no external criteria for validation. Hereunder are some important points to note in questionnaire design (Moser and Kalton, Chapter 13).

3.2 Questions should be relevant to survey objectives

Avoid covering too much or trying to collect information other than specific data needed. Lengthy questionnaire is always undesirable since extra length implies extra cost. In addition, respondents may become impatient and the reliability of their answers would be affected. Therefore, marginal questions should not be included.

3.3 Wordings must be clear, specific, precise and unambiguous

Vague words may lead to misunderstanding and produce invalid answers.

e.g. "How old are you?"

- some respondents may give answers in lunar year, some may round their age to next whole number while others to the last.
- better ask date (or year) of birth, or specify the way of calculating 'age'.

"*Don't* you think alternative X *is not* a good solution to problem Y?

- double-negative question, may confuse respondents.
- better ask "Do you think alternative X is a good solution to problem Y?"

"Have you seen the Dean of Students $\cdots$?"

- for a 'Yes' answer, its meaning may be
  (a) I have seen him (with my eyes),
  (b) I have consulted him or
  (c) I have met him, etc.
- better ask "Have you consulted $\cdots$?" if this is the subject of interest.

It is also important to ensure that the wordings used in a questionnaire mean the same thing to all respondents.

3.4 Employ appropriate language

Suitable language, including jargons and slangs, which is understood and accepted by respondents should be employed.

3.5 Be cautious when adapting questionnaires

When questionnaires or individual questions/scales in a *foreign language* are adapted for a local survey, be sure that the questionnaire items are compatible with local culture, otherwise appropriate modifications should be made. When questions in language A have been translated to language B, a helpful way to determine whether proper translation has been done is to ask another translator to translate them back to language A and see whether the two versions in language A differ significantly in meaning. If the difference is significant, improve the translation and repeat the procedure.

3.6 Memory error is an important source of inaccurate reporting

It is very difficult for someone to recall facts and events after a substantial lapse of time. Therefore a question should not relate to the distant past or a lengthy period of time. If unavoidable, respondents should be requested to refer to records.

3.7 Screening questions may be necessary

Some questions should be preceded by a screening question (i.e. filtering question).

e.g. "Does your part-time job have any adverse effect on your academic performance?"

- a 'No' answer given by the respondent may stand for
  (a) no part-time job, or
  (b) no adverse effect.

- this question should be preceded by "Do you have a part-time job?" Only those who answer 'Yes' to this question are required to answer the stated question.

3.8 Avoid leading questions

Some questions which by their content, structure or wording lead the respondent in the direction of a certain answer.

e.g. "You don't like ···, do you?"

- usually lead to negative answers.

"It is the government's responsibility to ···, don't you agree?"

- usually leads to affirmative answers.

"Do you watch any T.V. programmes (in the evening), such as A and B?"

- usually the answers are the named programmes.

3.9 Avoid loaded questions

There are ways in which a question may be loaded in favour of a particular response(s) (Smith, p.180).

e.g. "Do you accept *reasonable* fare increases for better ··· services?"

- the catch word reasonable increases the affirmative responses.

"*Confucius* says ···, do you agree?"

- prestige/famous names will favour a statement or an idea.

"As *everybody* knows ···"

"According to *the law* ···"

- citation of the status quo, social desirability, stereotypes etc. will get higher approval.

Such ways of asking questions are not recommended.

A questiona with an innocent statement may become loaded if the options for the answer are at varying degree of complexity. In this case, the question is often loaded against those options which are more

difficult to understand.

3.10 Avoid composite questions

It is important to avoid composite questions (or 'double barrelled' questions) where actually two or more questions have been combined into one.

e.g. "Do you plan to leave your job and look for another one within the next one month?"

- whereas one answer is expected for the question the respondent may actually want to answer 'Yes' to the first part of the question and 'No' to the second part.

e.g. "Do you have confidence on Hong Kong's *stability* and *prosperity* ···?"

- not unidimensional.
- better divide into two questions; one on stability, another on prosperity.

3.11 Use proper response categories

Closed questions (i.e. those with fixed alternatives as answers) are widely employed in large scale surveys because of their relative efficiency for responding, coding and analysis.

Response alternatives for closed questions should always be mutually exclusive and exhaustive. For opinion/attitude questions, response alternatives should preferably be symmetrical, with a middle or neutral alternative.

e.g. "How good is the ventilation of your flat?

(very good
(good
(acceptable
(poor "

- options not symmetrical, will produce more positive responses.
- better use

  (very good
  (good
  (average
  (poor
  (very poor

"How would you rate someone's performance ···?

(excellent
(good
(satisfactory
(unsatisfactory
(poor "

- not monotonic
- better use

  (very good
  (good
  (average
  (poor
  (very poor

"Are you satisfied with your present job?

(very satisfied
(satisfied, but feel pressure
(unsatisfied, but no pressure
(very unsatisfied"

- combining two scales.
- better divide into two questions; one on job satisfaction, another on work pressure.

On most occasions, alternatives such as 'not applicable', and 'no answer' (i.e. not willing to answer) should be added. Besides, alternatives such as 'no preference', 'don't know', 'no opinion', 'undecided', 'does not matter' and 'it depends' should be included as necessary, depending on the nature of the question. In deciding on the alternatives to be adopted the questionnaire designer should have a clear idea of what the selected ones refer to in the context of the question.

3.12 Good format and layout are important

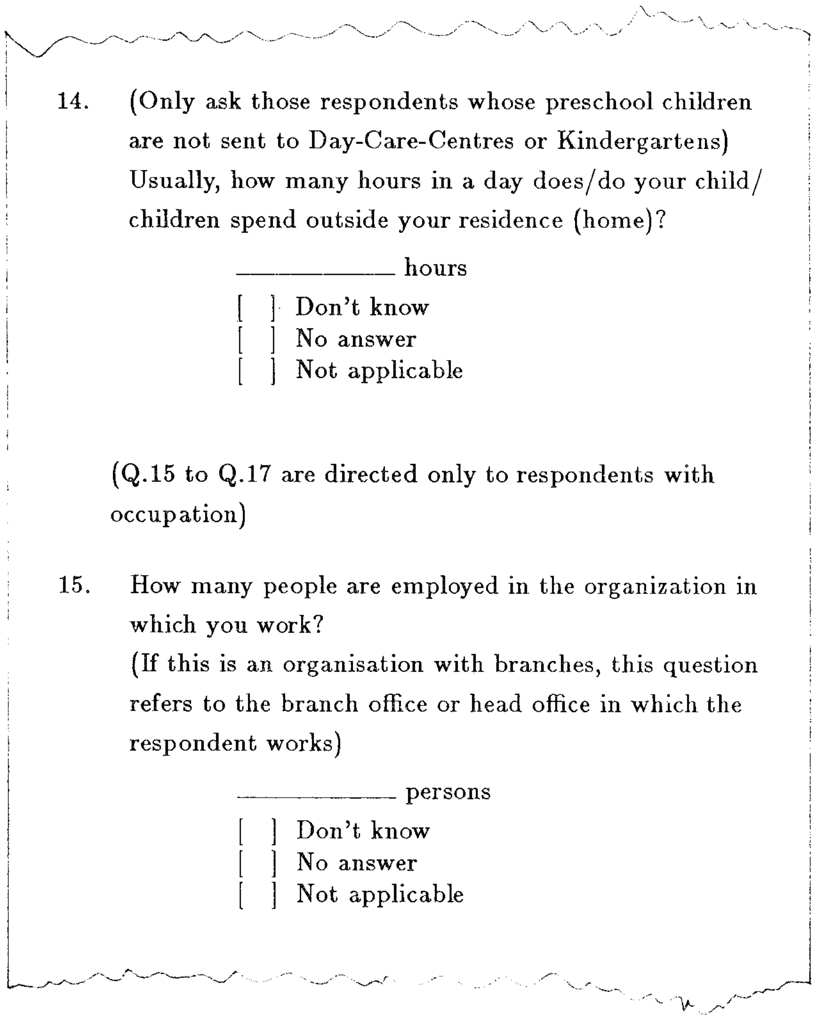
Questions should be arranged in proper order, divided into sections and numbered. There are no common principles of ordering, but note that sensitive and embarrassing questions, in case that they are unavoidable, should better be put at the end of the questionnaire. Good printing, optional wording and clear instructions (see Example A below), when/where to skip questions (see Example B below) etc., also help to eliminate mistakes in interviews and responding.

3.13 Pretest should be conducted

Questionnaires should be tested before they are put to use in surveys. Therefore, pretest should be carried out prior to the actual fieldwork, to find out whether the questions being asked serve the survey purpose(s), have the correct wording, response categories and ordering, and are appropriate to be asked of the target respondents.

References for this chapter can also be found in Babbie, Chapter 7.

Example A (Optional wordings and instructions in brackets)



14. (Only ask those respondents whose preschool children are not sent to Day-Care-Centres or Kindergartens) Usually, how many hours in a day does/do your child/children spend outside your residence (home)?

    ____________ hours

    [ ] Don't know
    [ ] No answer
    [ ] Not applicable

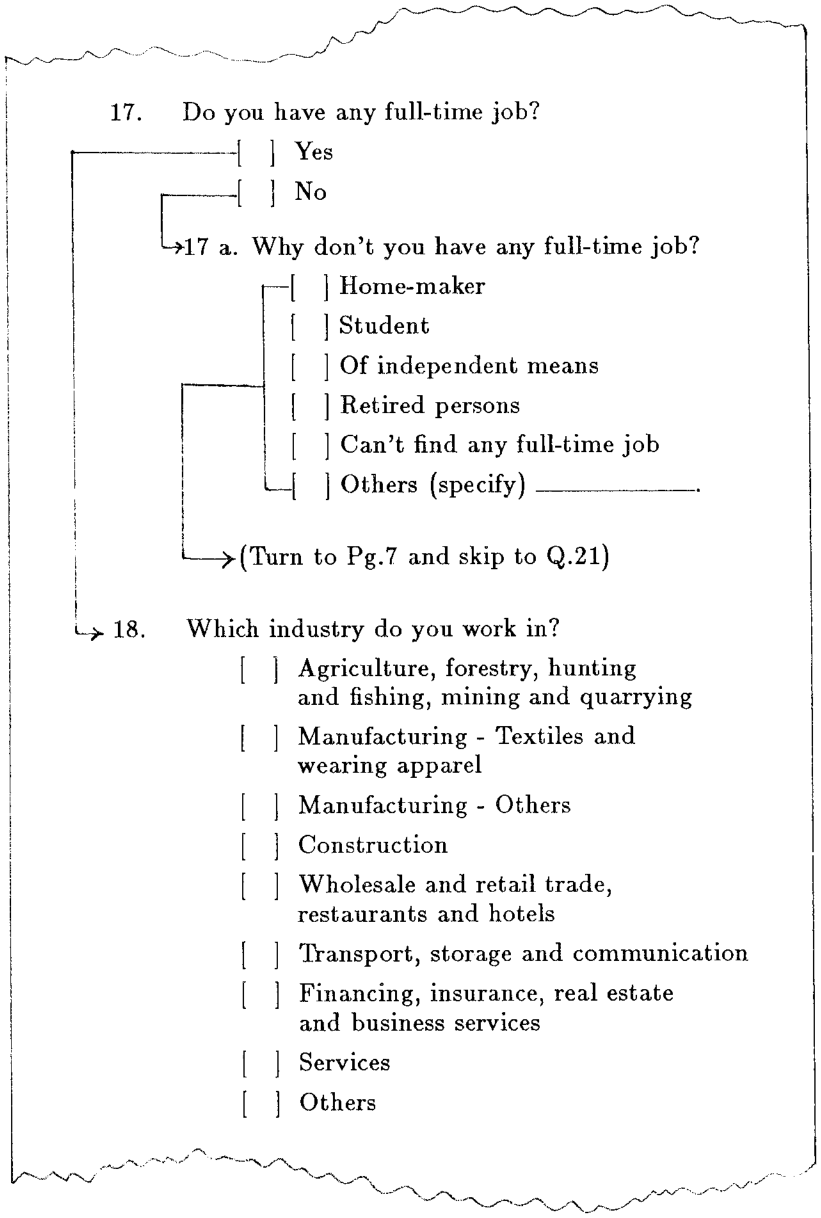
(Q.15 to Q.17 are directed only to respondents with occupation)

15. How many people are employed in the organization in which you work?
    (If this is an organisation with branches, this question refers to the branch office or head office in which the respondent works)

    ____________ persons

    [ ] Don't know
    [ ] No answer
    [ ] Not applicable

Example B (Flow-chart instructions for skipping questions)



17. Do you have any full-time job?
    [ ] Yes
    [ ] No

17 a. Why don't you have any full-time job?
    [ ] Home-maker
    [ ] Student
    [ ] Of independent means
    [ ] Retired persons
    [ ] Can't find any full-time job
    [ ] Others (specify) ____________.

(Turn to Pg.7 and skip to Q.21)

18. Which industry do you work in?
    [ ] Agriculture, forestry, hunting and fishing, mining and quarrying
    [ ] Manufacturing - Textiles and wearing apparel
    [ ] Manufacturing - Others
    [ ] Construction
    [ ] Wholesale and retail trade, restaurants and hotels
    [ ] Transport, storage and communication
    [ ] Financing, insurance, real estate and business services
    [ ] Services
    [ ] Others

CHAPTER 4
COLLECTION OF DATA

4.1 The pilot survey

The mode of data collection, such as the use of personal interviews, telephone interviews, self-administered questionnaires or postal questionnaires should be carefully selected by considering the respondents' willingness to co-operate, the degree of complexity of the subject of enquiry and other relevant factors.

Various arrangements related to fieldwork such as the proper allocation of workload amongst interviewers and the provision of adequate transport facilities, should be carefully planned to ensure smooth and efficient operation.

Fieldwork procedures should be thoroughly tested before implementation. A testing of the questionnaire and field procedures is the only way of finding out if everything works, especially if a survey employs a new procedure or a new set of questions. Since it is rarely possible to foresee all the possible misunderstandings or biasing effects of different questions and procedures, it is vital for a well-designed survey plan to include provisions for a pilot survey. Adjustments, where appropriate, will be made based on observations made and experience gained during the pilot survey.

4.2 Training interviewers

The interviewers should be carefully briefed on the concepts and definitions of terms used in the survey and properly trained in procedures before they start work.

The training may take the form of self-study, classroom training, or both. Training should stress on good interviewing techniques in making initial contacts, in conducting interviews in a professional manner and in avoiding influencing or biasing responses. The training generally involves practice interviews to familiarize the interviewers with the variety of situations they are likely to encounter in the survey. Survey materials must be prepared and issued to each interviewer, including sample copies of the questionnaire, a reference manual, information about the identification and location of the sampled units, and any cards or pictures to be shown to the respondent.

Interviewers should ensure that respondents understand the questions. Probing should be done only where necessary and should not be overdone in order to avoid exercising undue influence on respondents.

4.3 Minimizing non-response

A variety of procedures should be used in an attempt to minimize the number of refusals. Before conducting the interview, it is advisable for the survey organization to send an advance letter to the sampled units explaining the purpose of the survey and the fact that an interviewer will be calling soon. For large scale surveys, general publicity is important.

With interview surveys the interviewers are carefully trained on how to avoid refusals. It is often advisable to return to conduct an interview at a time more convenient to the respondent immediately on discovering that the first contact is not occurring at a convenient time (e.g. when a household is having dinner).

Attempts to persuade the sampled units of the value of the survey are, often supported by reference to a prestigious sponsor. Good sponsorship is likely to be particularly effective with a postal survey. Assurances of anonymity and confidentiality are generally provided to eliminate any fears the respondents may have about the use of their responses. Questionnaires are usually organized to start with simple, non-threatening questions to avoid the risk that the respondent will terminate the interview when immediately faced with an embarrassing question.

Not-at-homes in interview surveys are treated by callbacks. In face-to-face surveys, interviewers are commonly instructed to make at least 3 or 4 callbacks if unable to contact a respondent, with the callbacks having to be made on different days and at different times of day, including some evening calls. The interviewers are also encouraged to make additional calls if at all possible. Appointments can be useful in increasing the chance of contacting a respondent at a subsequent call.

In telephone surveys, callbacks are much more readily accomplished. The number of callbacks made is generally much larger than in face-to-face surveys.

In mail surveys the comparable procedure to callback is the follow-up, i.e. sending out further correspondence to urge those who have not replied. Often, a combination of callback methods may be made, for example, if respondents fail to reply by mail, an interviewer may visit them in person.

In general, every effort should be made by the survey-taker to minimize non-contacts with respondents or refusals to respond. Adequate publicity of the survey, proper identification of the surveying organization and interviewers, and giving advance notice to sampled respondents should help. Arrangements should also be made to follow up non-contact respondents and to persuade unco-operative respondents to participate.

4.4 Collecting observable characteristics of non-respondents

In the event that non-response is not negligible despite the huge effort to minimizing it, some proper methods should be devised to assist in the interpretation of survey results. For this purpose, it will be useful to obtain some observable characteristics of the non-respondents.

One type of non-response adjustment depends on information in the sample which is available for both respondents and non-respondents. For example, a sample is divided into geographical regions according to whether the sampled element is situated in a rural, suburban or metropolitan location. With an EPSEM (equal probability of selection of every member) sample, adjustments for variation in non-response rates across the resulting regions can be made by assigning weights of $n(h)/r(h)$ to the respondents in region h, where $n(h)$ is the total sample size selected, and $r(h)$ is the achieved sample size of respondents, in that region. These adjustments make the respondent sample distribution conform to the total sample distribution across the regions, with the respondents in a region being weighted up to represent the non-respondents in that region.

4.5 Quality control of collected data

Proper control of the progress of the survey should always be exercised. Last minute rush near the deadline causing deterioration in fieldwork quality should be avoided. Of equal importance is controlling the quality of individual interviews. This is normally done by having supervisory personnel to reinterview a subsample, and implementing office editing procedures to check for omissions or obvious mistakes in the data.

When the interviews have been completed and the questionnaires filled out, coding of questionnaire items which are not already precorded has to be done. Occupation and industry categorizations are typical examples of fairly complex coding for which quality control must be carefully exercised. This applies similarly to the coding of open-ended questions.

Data transcription and entry operations are subject to human errors and must be rigorously controlled through verification processes, either on a sample basis or 100 percent basis. Once a computer file has been generated, additional computer editing, as distinct from clerical editing of the data, can be accomplished to alter inconsistent or impossible entries (for example, a six-year old person who is reported to be a grandfather). Such a process of making alterations according to prescribed rules is known as *imputation.*

4.6 Guarantee of confidentiality

The privacy of the information supplied by survey respondents is of prime concern to all reputable survey organizations. The interviewers should clearly explain to the respondents that it is not the intention of the survey to describe the particular individuals in the sample. The survey is to obtain a statistical profile of the population. Individual respondents are therefore never identified and the survey results are always presented in the form of aggregate summaries, such as statistical tables and charts. With such assurance, respondents would probably be more willing to participate and give true answers.

In Hong Kong, the Census and Statistics Ordinance guarantees the confidentiality of data collected by the Census and Statistics Department. A number of professional organizations that rely on survey methods also have codes of ethics that prescribe rules for keeping survey responses confidential. The recommended policy for survey organizations to safeguard such confidentiality includes:

1. Using code numbers for the identity of a respondent on a questionnaire.

2. Refusing to release details of survey respondents to anybody outside the survey organization, including clients.

3. Destroying questionnaires after the responses have been processed.

4. Omitting the names and addresses of survey respondents from computer tapes used for analysis.

5. Presenting statistical tabulations by broad enough categories that data relating to individual respondents cannot be revealed through intelligent deduction.

References for this chapter can be found in Moser and Kalton, Chapters 2, 7, 11, 12 and 15, and Babbie, Chapters 9, 10 and 12.

CHAPTER 5
ANALYSIS OF DATA

5.1 Varied methods of analysis

When data have been edited they are ready for analysis. Reduction of the amount of information in the data and compilation of data to make analysis feasible are the essential steps to enable statistical description and statistical inference which are the major parts of survey analysis. The range of statistical methods applicable in the analysis of survey data is too great to be summarized here. Methods of finding and measuring patterns of association, basic principles of estimation and hypothesis testing, etc. can be found in standard statistical text books such as Freund (1988).

Often the main purpose of sample surveys is to estimate certain population parameters and different formulas are required for different sample designs. Some fundamental formulas are given in the following section for ease of reference.

5.2 Estimations from simple random samples

The mean $\bar{X}$, total X_T and proportion P of the population are the usual parameters of interest. Any estimation of a parameter based on a sample can never be accurate. For any estimation, therefore, some indication of the precision of the estimate must be attached. Attaching the standard error of an estimate is a common practice but this only reflects the error of estimate arises from the random sampling variation that is present when n of the units selected in a sample are measured instead of observing all the N units in the population. Other sources of error include non-response, measurement error and errors introduced in editing, coding and tabulation of the results. (Cochran, Chapter 13)

For simple random samples, the most frequently used estimate for the population mean $\bar{X}$ is the sample mean $\bar{x}$ defined as

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

where $x_1, \cdots, x_n$ are sample observations. The standard error of $\bar{x}$ is

$$\text{s.e. } (\bar{x}) = \sqrt{(1-f)}\frac{S}{\sqrt{n}}$$

where $f = n/N$ is the *sampling fraction* and S^2 is the variance of the finite population defined as

$$S^2 = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^2}{N-1}. \qquad (5.1)$$

In practice, the factor $1-f = (N-n)/N$, called the *finite population correction*, can be ignored whenever the sampling fraction f does not exceed 5% and for many purposes even if it is as high as 10%. The effect of ignoring the correction is to overestimate the standard error of the estimate $\bar{x}$.

For sampling with replacement the population is regarded to be infinite, the finite population correction is not required in the formulas and f is taken to be zero.

The computation of the standard error involves the population variance S^2 which in practice is unknown and can be estimated by

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}. \qquad (5.2)$$

Consequently the standard error of $\bar{x}$ can be estimated by

$$s(\bar{x}) = \sqrt{1-f}\frac{s}{\sqrt{n}}.$$

This estimate is slightly biased but for most applications the bias is unimportant.

It is usually assumed that the estimate $\bar{x}$ is normally distributed about $\bar{X}$. Then the lower and upper confidence limits for $\bar{X}$ are respectively

$$\bar{x} - z \times s(\bar{x}) \qquad \text{and} \qquad \bar{x} + z \times s(\bar{x}),$$

where z is the value of the normal variate corresponding to the desired level of confidence. The most commonly used values are:

Level of confidence (%)	50	80	90	95	99
z	0.67	1.28	1.64	1.96	2.58

If the sample size is less than 50 the percentage points may be taken from Student's t table with $(n-1)$ degrees of freedom. The t distribution holds exactly only if the observations x_i are themselves normally distributed and N is infinite. Moderate departures from normality do not affect the confidence limits greatly. For small samples with very skewed distributions, special methods are needed.

The computation of standard error of $\bar{x}$ serves three purposes: (1) to compare the precision obtained by simple random sampling with that given by other methods of sampling, (2) to estimate the size of the sample needed in a survey that is being planned as described in Section 2.3, and (3) to estimate the precision of the estimates actually attained in a survey that has been completed.

In the case of purpose (2), the population variance S^2 is required but its estimate s^2 based on sample observations is not available yet. There exist different methods to obtain an estimate. If other surveys conducted elsewhere which have studied similar characteristics in similar populations, often the measures of variability from those surveys can be applied as an indication of S^2 for the present population. If a pilot study has been conducted prior to a major sample survey, its

results may give some indication of the value of S^2. A more reliable approach is to take a preliminary simple random sample of size n_1 in order to obtain an estimate of S^2 for the determination of n and then augment the sample with a further simple random sample of size $(n - n_1)$. Occasionally some knowledge of the structure of the population may throw light to the value of S^2. For example, if the variable of interest can be modelled by a Poisson distribution, any information about the possible value of $\bar{X}$ can be used to estimate S^2.

The commonly used simple random sample estimate of the population total X_T is

$$x_T = N\bar{x}$$

whose standard error is

$$\text{s.e. } (x_T) = N\sqrt{1-f}\frac{S}{\sqrt{n}}$$

which can be estimated by

$$s(x_T) = N\sqrt{1-f}\frac{s}{\sqrt{n}}$$

where S and s are defined by equations (5.1) and (5.2) respectively. Analogous to the estimation of $\bar{X}$ the finite population correction $(1-f)$ can be ignored when f is small, also the lower and upper confidence limits for X_T are

$$x_T - z \times s(x_T) \qquad \text{and} \qquad x_T + z \times s(x_T)$$

respectively.

The estimates $\bar{x}$ and x_T are unbiased for $\bar{X}$ and X_T. Their other properties and the detail theories can be found in Cochran, Chapter 2.

The sample proportion p is the estimate for estimating the population proportion P and the standard error of p is

$$\text{s.e. } (p) = \sqrt{1-f}\frac{S}{\sqrt{n}}$$

where S is now defined as

$$S^2 = \frac{N\,P(1-P)}{N-1}.$$

Since P is unknown, S^2 is also unknown. An unbiased estimate for S^2 is

$$s^2 = \frac{n\,p(1-p)}{n-1}$$

and the standard error of p is therefore estimated to be

$$s(p) = \sqrt{1-f}\frac{s}{\sqrt{n}}$$

$$= \sqrt{1-f}\sqrt{\frac{p(1-p)}{n-1}}.$$

When the finite population correction can be neglected when f is small, the expression is simplified to

$$s(p) = \sqrt{\frac{p(1-p)}{n-1}}.$$

When the sample size n is not too small, normal approximation can be used and the confidence limits for P are approximately

$$p - z \times s(p) - \frac{1}{2n} \qquad \text{and} \qquad p + z \times s(p) + \frac{1}{2n}. \qquad (5.3)$$

The last term $\frac{1}{2n}$ in equation (5.3) is a correction for continuity without which the normal approximation usually gives too narrow a confidence interval. Details about sampling proportions can be found in Cochran, Chapter 3.

5.3 Estimations from more elaborated sample designs

The formulas given in the previous section have to be modified for more elaborated sample structures. The modified forms for simple stratified sampling, simple cluster sampling and systematic sampling

are summarised below. For details and for further types of sample designs readers can refer to Cochran.

In the following listings, capitals denote population parameters and lower case letters denote sample values.

Simple stratified random sampling — obtained by taking a simple random sample from every stratum in the population. The following notation will be used:

L	number of strata in the population
N_h	hth stratum size therefore $N_1 + \cdots + N_L = N =$ population size
$W_h = \frac{N_h}{N}$	hth stratum weight (relative stratum size)
n_h	hth stratum sample size therefore $n_1 + \cdots + n_L = n =$ sample size
X_{hi}	the ith member in the hth stratum
x_{hi}	the ith selected member from the hth stratum
$\bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} X_{hi}$	the hth stratum population mean
$\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}$	the hth stratum sample mean
$S_h^2 = \frac{1}{N_h-1}\sum_{i=1}^{N_h}(X_{hi} - \bar{X}_h)^2$	the hth stratum variance
$s_h^2 = \frac{1}{n_h-1}\sum_{i=1}^{n_h}(x_{hi} - \bar{x}_h)^2$	the hth stratum sample variance
P_h	the hth stratum population proportion
p_h	the hth stratum sample proportion

Denote the estimate for $\bar{X}$ in a simple stratified random sampling by $\bar{x}_{\mathrm{st}}$ the formulas are:

$$\bar{x}_{\mathrm{st}} = \sum_{h=1}^{L} W_h \, \bar{x}_h$$

$$\text{s.e. } (\bar{x}_{\mathrm{st}}) = \sqrt{\sum_{h=1}^{L} W_h^2 (1 - f_h) \frac{S_h^2}{n_h}}$$

where $$f_h = \frac{n_h}{N_h}$$

$$s(\bar{x}_{\mathrm{st}}) = \sqrt{\sum_{h=1}^{L} W_h^2 (1 - f_h) \frac{s_h^2}{n_h}}.$$

Denote the estimate for X_T in a simple stratified random sampling by x_T, the formulas are:

$$x_T = N\bar{x}_{\mathrm{st}}$$

$$\text{s.e. } (N\bar{x}_{\mathrm{st}}) = N \times \text{s.e. } (\bar{x}_{\mathrm{st}})$$

$$s(N\bar{x}_{\mathrm{st}}) = N \times s(\bar{x}_{\mathrm{st}}).$$

Denote the estimate for P in a simple stratified random sampling by p_{st}, the formulas are:

$$p_{\mathrm{st}} = \sum_{h=1}^{L} W_h \, p_h$$

$$\text{s.e. } (p_{\mathrm{st}}) = \sqrt{\sum_{h=1}^{L} W_h^2 \frac{N_h - n_h}{N_h - 1} \frac{P_h(1 - P_h)}{n_h}}$$

$$s(p_{\mathrm{st}}) = \sqrt{\sum_{h=1}^{L} W_h^2 \frac{N_h - n_h}{N_h - 1} \frac{p_h(1 - p_h)}{n_h}}.$$

Simple cluster sampling — obtained by taking a simple random sample of m clusters from a total of M from the population and including in the sample all members of the chosen clusters. The notation is:

M	number of clusters in the population
m	number of clusters in the sample
N_i	the ith cluster size
	therefore $N_1 + \cdots + N_M = N$
n_i	the ith selected cluster size
	therefore $n_1 + n_2 + \cdots + n_m = n$
X_{ij}	the jth member in the ith cluster
x_{ij}	the jth member in the ith selected cluster
$\bar{X}_i$	the ith cluster mean
$\bar{x}_i$	the ith selected cluster mean
X_{iT}	the ith cluster total
x_{iT}	the ith selected cluster total
P_i	the ith cluster proportion
p_i	the ith selected cluster proportion

There are three frequently used estimates for $\bar{X}$:

(a) The 'cluster sample ratio' denoted by $\bar{x}_{ca}$

$$\bar{x}_{ca} = \frac{\sum_{i=1}^{m} x_{iT}}{\sum_{i=1}^{m} n_i}$$

$$\text{s.e. } (\bar{x}_{ca}) \doteq \sqrt{\frac{(M - m)M}{(M - 1)m} \sum_{i=1}^{M} \left(\frac{N_i}{N}\right)^2 (\bar{X}_i - \bar{X})^2}$$

$$s(\bar{x}_{ca}) \doteq \sqrt{\frac{(M - m)M}{(m - 1)m} \sum_{i=1}^{m} \left(\frac{N_i}{N}\right)^2 (\bar{x}_i - \bar{x}_{ca})^2}.$$

If N is unknown, it can be replaced by an estimate Mn/m.

(b) The 'cluster sample total' denoted by $\bar{x}_{cb}$

$$\bar{x}_{cb} = \frac{M}{Nm}\sum_{i=1}^{m} x_{iT}$$

$$\text{s.e. } (\bar{x}_{cb}) = \sqrt{\frac{(M-m)M}{(M-1)mN^2}\sum_{i=1}^{M}(X_{iT} - \bar{X}_T)^2}$$

$$s(\bar{x}_{cb}) = \sqrt{\frac{(M-m)M}{(m-1)mN^2}\sum_{i=1}^{m}(x_{iT} - \bar{x}_{cb})^2}.$$

(c) The 'unweighted average of the chosen cluster means' denoted by $\bar{x}_{cc}$

$$\bar{x}_{cc} = \frac{1}{m}\sum_{i=1}^{m} \bar{x}_i$$

$$\text{s.e. } (\bar{x}_{cc}) = \sqrt{\frac{M-m}{mM(M-1)}\sum_{i=1}^{M}(\bar{X}_i - \bar{X}_c)^2}$$

where $$\bar{X}_c = \frac{1}{M}\sum_{i=1}^{M} \bar{X}_i$$

$$s(\bar{x}_{cc}) = \sqrt{\frac{M-m}{mM(m-1)}\sum_{i=1}^{m}(\bar{x}_i - \bar{x}_{cc})^2}.$$

To compare the three estimates $\bar{x}_{ca}$, $\bar{x}_{cb}$, $\bar{x}_{cc}$ for the population $\bar{X}$, some properties are summarised in the following table:

	Estimate		
	$\bar{x}_{ca}$	$\bar{x}_{cb}$	$\bar{x}_{cc}$
Unbiasedness	biased, unimportant only when m is large	unbiased	biased, not too serious when N_i do not vary too much unless the cluster means and cluster sizes are highly correlated
Standard error	depends on the variation between cluster means: smaller the variation, smaller the standard error	tend to be larger than that of $\bar{x}_{ca}$ unless N_i do not vary too much	the expected mean square error is more relevant then the standard error unless N_i do not vary too much
Estimate for X_T	$N\bar{x}_{ca}$	$N\bar{x}_{cb}$ $= \frac{M}{m}\sum_{i=1}^{m} x_{iT}$	$N\bar{x}_{cc}$
Relative merits	more efficient	unbiased, does not need to know N for estimating X_T	simple and quick

Denote the estimate for P in a simple cluster random sample by p_c, some formulas are:

$$p_c = \frac{M}{Nm}\sum_{i=1}^{m} N_i\, p_i$$

$$\text{s.e. } (p_c) = \sqrt{\frac{M-m}{mM(M-1)\bar{N}^2}\sum_{i=1}^{M}(N_i P_i - \bar{N}\,P)^2}$$

where $\bar{N} = \frac{1}{M}(N_1 + \cdots + N_M)$

$$s(p_c) = \sqrt{\frac{M-m}{mM(m-1)} \sum_{i=1}^{m} \left(\frac{M}{N} N_i p_i - p_c \right)^2}.$$

Systematic sampling — sample members are chosen at regular intervals from a complete list of the population members.

A systematic sample can be viewed as a cluster sample where the population is divided into M clusters. The members belonging to each of the M clusters are as follows:

1st cluster	$X_1,$	$X_{M+1},$	$X_{2M+1},$	$\cdots$
2nd cluster	$X_2,$	$X_{M+2},$	$X_{2M+2},$	$\cdots$
$\vdots$				
Mth cluster	$X_M,$	$X_{2M},$	$X_{3M},$	$\cdots.$

Each cluster is of sizes n or $n+1$ and the sample consists of one cluster ($m = 1$) chosen randomly from the M clusters. If the population list is arranged in a random order, a systematic sample can be treated as a simple random sample. For more details Cochran, Chapter 8 may be referred to.

The formulas listed are the very basic ones for the simplest types of sample design. The more complex situations will inevitably involve multistage sampling, concomitant variables and cost factors, either individually or in combination. For technical details, readers should refer to books such as Cochran or consult survey experts.

5.4 Be cautious with small subgroups

Surveys are very often multipurpose, where a substantial number of variables are studied simultaneously. Moreover, population values have to be estimated not only for the total population but also for a wide range of subgroups, perhaps for people in different age groups, at different educational levels, different geographical regions and so on.

A sample size sufficiently large to provide reliable estimates at the aggregate level may be inadequate to support subgroup analysis if the subgrouping is too refined. This is because with a refined subgrouping scheme, the 'effective sample size' that corresponds to any of the subgroups in question would be quite small such that the sampling error of the variable (hence the margin of error of the estimate) would be large. Due deliberations should therefore be made to decide on an appropriate subgrouping scheme in analyzing the survey results.

Larger samples permit finer divisions of the sample for subgroup analysis, and the choice of sample size often depends on an assessment of the costs of increasing the sample compared with the possible benefits of more detail analyses.

5.5 Importance of a sufficiently high response rate

For a given survey population, the response rate for the survey is defined as the ratio of the number of questionnaires completed for eligible elements to the number of eligible elements in the sample. According to this definition, ineligible elements such as blanks and foreign elements (see §2.5) should be excluded from both the numerator and denominator in calculating the rate.

A low response rate may do even more damage in rendering a survey's results questionable than a small sample, since it is very difficult to infer the characteristics of the population represented by the non-respondents. Therefore, a sufficiently high response rate has to be achieved in the survey before its results can provide valid estimates of the population under study. For a survey which has a low response rate, the survey results can be very misleading.

The failure to collect the survey data from some sampled elements, or non-response, is a major survey problem. The cause of concern about non-response is the risk that non-respondents may differ from respondents with regard to the survey variables, in which case the survey estimates based on the respondents alone will be biased estimates of the overall population parameters. The following example illustrates the concept of non-response bias.

Consider a situation in which the population is divided into two groups — those who respond and those who do not. Suppose that the aim of a survey is to determine M, the total population mean. This mean my be expressed as:

$$M = W(r)M(r) + W(m)M(m)$$

where $M(r)$ and $M(m)$ are the means for the response and non-response group (the r stands for respondents and m for missing), and $W(r)$ and $W(m)$ are the proportions of the population in these two groups $[W(r) + W(m) = 1]$. Since the survey fails to collect data for the non-respondents, it only produces the estimate $M(r)$.

The difference between $M(r)$ and the population mean M is:

$$M(r) - M = W(m)[M(r) - M(m)].$$

This difference, which is the *non-response bias* arising from using the respondent mean in place of the overall mean, is seen to depend on two factors: $W(m)$, the proportion of non-respondents in the population; and $[M(r) - M(m)]$, the difference between the means of respondents and non-respondents.

If the respondent and non-respondent means are equal or very close, there would be no, or very small, non-response bias. In practice, however, it is inappropriate to assume that the missing responses are missing at random. It is likely that the non-respondents have characteristics different from respondents. Therefore, the only way to make

sure that non-response bias is not sizable is to keep the non-response group sufficiently small (i.e. small $W(m)$). Weighting of a sample to a known population distribution will adjust for non-response as well as non-coverage (the failure of some elements in the survey population to be included on the sampling frame). For example, the age distribution of the population is known from a recent population census. If, in a survey, the non-response rate is higher among young people, or if more of them are missing from the sampling frame, then weighting the sample to make it conform to the known age distribution will compensate for these factors. However, to the extent that there are differences in the survey variables between the respondents and non-respondents within each age group, some non-response bias will remain.

Often there are also inappropriate gaps in the data records for the respondents. The respondents may not know the answers to certain questions, or they may refuse to answer some questions because they find them sensitive, embarrassing, or they consider them irrelevant to the perceived survey objectives. Also, an interviewer may incorrectly skip over a question or fail to record an answer. Even when an answer is recorded on the questionnaire, it may be rejected during editing prior to analysis because it is inconsistent with other answers. These inappropriate gaps in the data records for the respondents are called item non-response.

Item non-response rate is defined as the number of eligible elements failing to provide an answer to the item divided by the total number of eligible elements in the sample. Corresponding to the weighting adjustments for non-response, various imputation methods have been devised to try to compensate for the bias of item non-response. A major benefit of imputation is that a data set for the respondents with no missing values is constructed, which greatly facilitates survey analyses.

One must be prepared, in the event of a very low response rate, to accept that the survey has been a failure and hold back the results. Publicising results regardless of the response rate is not a responsible move. Simply 'cautioning' the readers of survey results is not enough.

References for this chapter can also be found in Babbie, Chapter 16.

CHAPTER 6
DISSEMINATION OF SURVEY RESULTS

A complete survey report should contain details about the different aspects of the survey, in particular details on the population covered, sample design, margin of error, response rate and likely sources of non-sampling error. Specimens of the questionnaire should be attached. This is particularly important with surveys of opinions. Technical details should be included, preferrably in appendices. Investigators should bear in mind that style, language and presentation of survey reports depend on their readership.

When the results of a survey with public implications are released to the mass media, the survey-taker should supply sufficient details on survey methodology in addition to survey findings, so that the mass media may report both, otherwise the general public will have no basis to assess the reliability of those findings.

If there are biases in the survey design — for good reason or otherwise — the survey-taker should make these design flaws and their consequences known to the general public. The possible biases may, for example, include sample selection bias, weighting bias, question bias and question order effect known to the survey-taker. If there are any of these biases, the limitation of any survey conclusions should be clearly spelled out.

For example, if non-probability samples have inevitably to be used, the likely bias of the method should be properly highlighted and there should be cautions against generalizations from the sample to some larger population.

Many survey reports or press stories about surveys tend to skip any mention about methodology. Probably the authors think that the

the readers are not interested or cannot understand. This may well be true but then the readers can be misled. It is hoped that the authors can present the methodology in a simple, easy-to-understand and non-boring manner, and that the general public will become better educated statistically to understand commonly used methodology.

It is recommended that the following information should not be ignored in the dissemination of opinion survey results:

1. Sponsorship of the survey,
2. the time period of data collection,
3. sampling method,
4. mode of data collection (see §4.1),
5. wording of questions,
6. population covered,
7. sample size and response rate,
8. sample sizes and response rates for subgroup analysis,
9. margin of error (or confidence interval) and if possible,
10. likely sources of non-sampling error.

References for this chapter can be found in Babbie, Chapters 18, 19 and 20, and in Hage *et al*, pages 104–105.

REFERENCES

Babbie, E. (1990). *Survey Research Methods.* 2nd ed. Wadsworth.

Cochran, W.G. (1977). *Sampling Techniques,* 3rd ed. Wiley and Son.

Freund, J.E. (1988). *Modern Elementary Statistics.* 7th ed. Prentice-Hall.

Hage, G., Dennis, E., Ismach, A. & Hartgen, S. (1976). *New Strategies for Public Affairs Reporting.* Prentice Hall.

Kalton G. (1983). *Introduction to Survey Sampling.* Sage Publication.

Moser C.A. & Kalton G. (1971). *Survey Methods in Social Investigation.* Heinemann Educational Book Ltd. E.L.B.S. edition of 2nd ed.

Smith, H.W. (1975). *Strategies of Social Research — the Methodological Imagination.* Prentice Hall.