
UNIT 21 POLICY ANALYSIS: METHODS AND TECHNIQUES–II

Structure

21.0 Learning Outcome

21.1 Introduction

21.2 Survey Research

21.2.1 Types of Survey

21.2.2 Survey Methods

21.2.3 Questionnaire-Design

21.2.4 Pretesting a Questionnaire

21.2.5 Sampling Methods

21.2.6 Sample Size Determination

21.3 Computer Simulation

21.3.1 A Simple Simulation Model

21.3.2 Reasons for Using Simulation

21.3.3 Limitations of Simulation

21.3.4 Computer Simulation

21.3.5 Some Applications of Computer Simulations in Public Policy

21.4 Conclusion

21.5 Key Concepts

21.6 References and Further Reading

21.7 Activities

21.0 LEARNING OUTCOME

After studying this Unit, you should be able to:

- Understand the concepts of survey research and computer simulation;
- Discuss the methods of collecting data through survey research;
- Explain the strengths and weaknesses of different survey methods;
- Describe the ways of conducting computer simulations; and
- Highlight the advantages and disadvantages of computer simulation.

21.1 INTRODUCTION

The present day society is an information society. This society needs information on needs, preferences, and behaviour. The government, business, and social institutions also need information so that they can meet the preferences and expectations of individuals. In this context, information, as processed data, is essential as an aid to decision-making. The data can be either gathered through surveys or generated through computer simulation.

This Unit provides clear understanding on how to collect data through surveys and computer simulation. For this purpose, choose a method on the basis of the nature of population and its needs, and assess the limitations of data collection. The computer simulations are useful to understand the system that helps in decision-making process. You may prefer to use computer simulations, particularly when the methods of mathematical optimisation procedures are complex. You will learn a more details on these topics in the rest of this Unit.

21.2 SURVEY RESEARCH

Survey research is the method of collecting data from a sample of respondents, that is, representative of a target population. It is one of the most common types of social research for collecting quantitative data. The survey is intended to represent characteristics or perceptions of a target population. It consists of all objects, individuals, or social units. For example, to know the environmental effects of an industry in a particular location, the target population may consist of all the individuals living in and around that industrial area. Similarly, to know the people's opinion on construction of an irrigation project the target population may consist of the potential beneficiaries of the proposed project and those who are likely to lose land or get displaced. Therefore, the target population must be clearly defined.

For a variety of reasons, it may not be possible for us to elicit the required information from the whole population. In such cases, we may use a sample that represents the target population. A sample is a fraction of population selected to elicit the information about the population as a whole. For example, a sample of prospective voters is questioned in advance to determine the public opinion on the candidates or parties contesting the election. Similarly, the government may seek the opinion of people through a sample survey for evaluating the present or proposed legislation.

21.2.1 Types of Survey

Normally, before conducting a survey one should decide as to how the results would be used. That is, whether the survey collects data at a single point of time or over a period of time. The first type of survey is called cross-sectional survey, and the second one longitudinal survey,

i) Cross-Sectional Survey

Cross-sectional survey is used to gather data on a single population or several populations at a single point of time. The characteristics of these populations can be examined on the basis of a number of factors. For example, if a city council is bringing a new legislation on road safety, it may want to know the opinion of people before the new policy is implemented. Therefore, it may conduct a cross-sectional survey among the city citizens. The population groups for this survey may consist of heavy vehicle drivers, light vehicle drivers, cycle and auto-rickshaw drivers, and pedestrians. The factors could be perceptions on speed limits, type of punishment for road safety rule violators, and how the money generated through fines should be spent. One may also divide the population on the basis of their income, educational qualifications, age, etc. and analyse their opinions.

ii) Longitudinal Survey

Longitudinal survey gathers data on a particular population over a period of time to permit comparisons. For example, if the government is introducing a new policy on electricity billing, bill distribution and collection in a city, it may conduct the opinion survey of its people before and after introduction of new policy. The longitudinal survey can further be divided into trend studies, cohort studies, and panel studies.

● Trend Studies

The trend studies aim at collection of data at regular intervals through repeated surveys, such as, surveys on unemployment, poverty, consumer prices, etc. While samples are of the same population they are typically not composed of the same people. Trend studies, may be conducted over a long period of time. The researcher can combine data from several studies of the same population in order to show the trend(s).

● Cohort Studies

Cohort studies also focus on a particular population or sub-groups of a population. They are sampled and studied more than once, but these studies have a different focus. They study a group, which has experience of an event or a programme or scheme. For example, a sample of students who have completed master's degree in Public Policy in 2006 can form a cohort and could be questioned regarding the utility of the programme in their respective professions. Five years later a different sample of the students who have completed their master's degree in Public Policy in 2006 can be studied to see any changes in their perceptions. Therefore, a cohort study would sample the same population at different points of time. Suppose, if you study the students who have completed their master's degree in 2008 then it becomes a trend study and not cohort study.

● Panel Studies

Panel studies are similar to cohort studies, except that the same individuals are interviewed in each period, whereas in cohort studies only a random sample of the same group (year, age, income, etc.) is studied. In Panel studies the same set of individuals serve the control function over a period of time, whereas in Cohort analysis random sampling is assumed to serve as control function, assuring similarity of a Cohort by age, education, gender, etc. as it is measured in successive time periods. For example, select a sample of voters and ask them about their affiliation to political parties. Every year thereafter, the researcher would contact the same voters and ask them similar questions, and ask them the reasons for any changes in their preferences.

The Cohort studies measure net changes over time, whereas the Panel studies measure gross changes over time. For example, many voters switch their party identification in a given period of time and it does not take into account the new voters to measure the net effect. Panel studies can yield specific information and extremely useful explanations. But they are difficult to conduct. They tend to be expensive; they take lot of time and suffer from attrition rates (attrition occurs when people dropout of the Panel Study).

21.2.2 Survey Methods

Broadly speaking, there are four methods of conducting survey for collecting data: (a) in-person interview, (b) telephone survey, (c) mail survey, and (d) internet survey. Each of these methods has particular advantages and disadvantages. The choice of a particular survey method for data collection should be justified by its advantages and disadvantages in the context of the purpose for which a survey is being conducted.

i) **In-person Interview:** Basically there are three approaches to collect data through in-person interviewing: (i) informal conversational interview, (ii) general interview-guide interview, and (iii) standardised format interview.

The *informal conversational interview* relies on spontaneous questions arising from the natural flow of conversation. During this kind of interview, the respondents may not even know or realise that they are being interviewed. The advantage of this method is that it allows the people to respond to individual and situational differences. The questions can be personalised in Focus and

the communication would be non-threatening. The disadvantage of the informal interview is that it requires a great deal of effort and time to collect the information systematically. There is always scope to miss some important data since during the conversation uniform questions cannot be covered. Also, the quality of the data depends on the skills of the individual investigator.

Before interviewing the respondent, an interview-guide is prepared that contains issues or questions to be asked during the interview. It ensures that the data is collected on the same issues. The guide provides basic topics or subject areas within which the investigator is free to probe and obtain information on the subject. The main advantage of the interview-guide method is that it ensures that the investigator uses the limited time available in the best possible manner. The interview will be more systematic and comprehensive with focused discussion. This method is especially useful for group discussions, where the interview-guide keeps the discussion focused, but at the same time allows the individual perspectives to be identified. There are some disadvantages of this method. Although, this method overcomes the deficiency of omitting important issues as in informal interviews, still it can inadvertently omit a few important issues. In this method, the investigator can sequence the questions in his own way or use his own words in putting questions, which may often be misunderstood by the respondents and lead to inconsistencies in responses. Also, the reliability of the data greatly depends on the skills of the investigator.

When the data needs to be compared and used for further statistical analysis, the data should be collected using a standardised *format*. In this method, each respondent is asked essentially the same questions. The format can contain open-ended questions or closed-ended questions or both. The questions are sequenced as the investigator asks them. The advantage of the standardised format interview is that it minimises investigator's bias by standardising the questions being asked of each respondent. The interview is systematic and data comparisons become easier. Also, questions will be highly focused and reduces the time duration of the interview. This method has one disadvantage. The method does not allow the interviewer to pursue issues that may emerge during the period of interview.

Generally, the in-person interviews are the most expensive form of survey research. However, the response rates are very high and the researcher obtains more detailed data through this method. Also, in-person interview methods need larger and highly trained persons to collect the data when compared with other methods. The investigator's personal bias also forms one of the drawbacks of this method. A number of sampling methods are used to draw samples from population under study. We will discuss these methods in section 21.2.4 of this Unit.

ii) **Telephone Surveys:** Telephone surveys offer a comparatively fast and low-cost alternative to in-person interviews and are particularly useful when the population is large and geographically dispersed. One of the drawbacks of this method is that it is limited to persons with telephones. Also, if the telephone book is used as the source of names, persons who have telephones but their names are not listed in the book are omitted from drawing samples for surveys. Another major disadvantage is that many people are switching over to mobile phones, which is presented in Unit 24.

To draw a sample of respondents, the investigator starts with a sampling frame containing phone numbers, chooses one unit from this frame, and conducts an interview over the telephone, either with a specific person at the number or anyone available at that number. This process is continued till all the sample units in the sampling frame are interviewed. Another technique called random-digit dialling permits telephone calls to be made within specific geographic areas without the need for the list of telephone numbers. However, the drawback of this method is that the interviewer dials a number not knowing whether there definitely exists a live connection at that number, or whether it is a business, hospital or household. It is always better to combine these two methods. First, the interviewer uses random digit dialling to prepare an initial list of random numbers. Using

a random number mechanism (we will discuss random numbers in section 21.2.5 of this Unit), telephone numbers are then taken from this list to produce a final set for the sample to be covered. The greatest advantage of telephone survey is that it is timely and costs less than in-person interviews. But the disadvantage is that the telephone survey requires several interviewers to make telephone calls if the survey is to be completed within a reasonable time period. These days most of the commercial firms use telephone to conduct surveys. Properly conducted telephone surveys yield high response rates. However, the interviewer should keep the number of questions to be asked to the minimum and simple. Researchers generally believe that they get the best response rate through in-person interviews. But studies show that the response rates have been declining particularly from urban areas and with high-income groups. Studies have shown that telephone interviews yield about as representative a sample as in-person interviews.

iii) **Mailed Survey:** In this method a pre-designed questionnaire is mailed to the respondents, who are expected to complete and return the questionnaire. The main advantage of mailed survey is that it is substantially less costly than both in-person and telephone surveys. In order to improve the response rate, the questionnaire is developed with closed-ended questions that make the data analysis relatively simple and quantifiable. Responses can be directly compared and easily aggregated. It is commonly observed that the response rates in mailed surveys are low. A high non-response rate sometimes results in non-representative sample. The non-response rates can be caused by many factors, which include non-availability of respondents, or refusal to participate in the survey. A variety of techniques can be used to improve the response rate. For example, a follow-up of the mail can improve the response rate. Also, the questionnaires that are brief are more likely to be completed. Some kind of attraction can be made to get the attention of the respondents to fill and return the questionnaire. For example, the questionnaire can have copy of an introductory letter from a prominent person, and respondents can be given token gifts. Another disadvantage of mailed questionnaires is that the respondents cannot fit their experience and views in the pre-determined categories. This can distort what the respondents really mean by limiting their choices. To partially overcome these difficulties, open-ended questions are often added to the questionnaires. Again open-ended questions are generally of limited value on mail survey because they depend entirely on the respondent to answer fully and do not provide an opportunity to probe or clarify unclear answers.

iv) **Internet Survey:** A more recent innovation in survey research methods is the use of Internet technology. There are two ways of contacting respondents through Internet. First, the researcher can mail the questionnaire using e-mail facility. In this method, you need to collect all the e-mail addresses to which you intend to mail the questionnaire. In the second method, the researcher simply posts the questionnaire in a web page with a request to the respondents to fill them. The respondents who access that web page download the questionnaire and send the filled-in questionnaire using either e-mail facility or postal mail. The respondents can also send their responses by simply filling the questionnaire on-line using the web page.

The Internet surveys substantially reduce the cost and time of the surveys. The major limitation of the Internet mails is that the population whom the Internet survey is aimed at should have Internet connectivity. Otherwise, there is a threat for sampling bias. If the target population consists of computer users, the sampling bias may be minimal. Therefore, the Internet surveys are ideal for collecting data from networked people and with e-mail addresses.

21.2.3 Questionnaire-Design

A questionnaire is a format containing questions arranged in a sequence to elicit information from the respondents. A questionnaire can be used for in-person interviews, or mailed surveys or Internet surveys. In a way, the questionnaire is the heart of the survey research. The researcher should be

clear in his/her mind what information is needed and from whom? This is crucial whether the questionnaire is administered by the interviewer or mailed or posted in a web page. Also, provide some introduction describing the nature of the questionnaire and the objectives of the survey. Mailed surveys and Internet surveys should have a brief attachment introducing the survey and requesting the cooperation of the respondent in filling the questionnaire.

Types of Questions: Each question or set of questions should be designed and carefully worded. They should have clear instructions on how the respondents have to fill the answers. For example, "What difficulties are you facing in your business? Check the reason that applies in your case.",

Closed-ended questions limit the respondent's answers to the survey. The respondents are allowed to choose from either a pre-existing set of dichotomous answers, such as, Yes/No, True/False, or multiple choice responses with an option for 'other' to be filled in or ranking scale response options (like very important, somewhat important, not too important, not at all important). The following are some examples of closed ended questions.

i) **Educational Qualifications:**

- a) Illiterate
- b) Below matriculation
- c) Matriculation and above, but below graduate
- d) Graduate and above

ii) **Which TV channel would you generally prefer:**

- a) Sports
- b) Entertainment/music/movies
- c) News
- d) Spiritual
- e) Other (please specify)

iii) **The most important problem that our country is facing in the current year is (rank the importance of each of the following reasons you perceive strongly):**

	I strongly agree	I somewhat agree	I have no opinion	I somewhat disagree	I strongly disagree
Poverty	()	()	()	()	()
Population	()	()	()	()	()
Corruption	()	()	()	()	()
Terrorism	()	()	()	()	()
Illiteracy	()	()	()	()	()
Employment	()	()	()	()	()

The advantage of closed-ended questions is that they are more easily analysed. The response to every answer can be quantified by assigning a pre-coded number so that statistical analysis becomes easier. Since closed-ended questions are more specific and convey similar meaning. They facilitate comparisons of responses. The time taken by a respondent or interviewer is limited. Because of this reason, the response rates are higher in surveys that use closed-ended questions than open-ended questions.

The advantage of open-ended questions is that they allow respondents to include more information, including feelings, attitudes and understanding of the subject. However, the response rates may not be encouraging when open-ended questions are used.

Wordings: The wording of survey questions is a tricky one. It is difficult to develop questions with proper words that convey same meaning between the researcher and the respondents. To write effectively communicating questions the researcher should keep in mind the following:

- Questions should be written in simple language and avoid complex or long-winding sentences.
- Questions should be kept short and simple.
- Questions should be specific and focused. Avoid general questions.
- Avoid questions that are sensitive and hurt the feelings of respondents.

Order of Questions: The sequence of questions is another important consideration in developing a questionnaire. The following points should be kept in view:

- Easier questions that the respondent feels that he/she can respond should be at the beginning. For example, you can put questions like name, age, qualifications, sex, etc. at the beginning of the questionnaire.
- Sensitive questions (if any) should be avoided at the beginning of the questionnaire.
- Do not put most important questions at the end of the questionnaire. Many respondents may not answer the questions appearing at the end because they are likely to develop fatigue by that time.
- The questions should be ordered in such a way that corresponds to the respondent's natural way of thinking.
- One should be careful while designing contingency questions like 'E your answer is yes to the previous question', etc.

21.2.4 Pre-testing a Questionnaire

Pre-testing of a questionnaire is essential to improve the effectiveness of a questionnaire for collecting data. Before sending the questionnaire to the respondent or the interviewer meets the respondents, the draft questionnaire should be tested for its effectiveness with a small group of people similar in characteristics to your ultimate sample. Pre-testing can help you determine the strengths and weaknesses of your survey concerning question format, wording, and order. There is no hard and fast rule to decide the number of respondents for pre-testing. However, it is ideal to have a pre-test sample of 10 and not more than 30.

21.2.5 Sampling Methods

Before conducting a survey, you must choose a representative sample of the population under study. Unless your target population is very small, it is not always possible to collect the data from all members of the population. A subset of population that contains all the characteristics of that population is called a sample. In that case, all the inferences drawn from the sample are probably applicable to the entire population. Therefore researchers generally survey a sample list of respondents from a population, which we call a sampling frame. There are a number of sampling methods, which can be classified into: (a) probability sampling methods, and (b) non-probability sampling methods. These are also sometimes referred as random sampling methods, and non-random sampling methods respectively.

Probability Sampling Methods: In the probability sampling methods each unit or item in the population has a chance of being included in the sample. This method of sampling reduces the sampling errors and normally we say that the sample is an unbiased sample. There are several methods of probability sampling. In this Unit, we will discuss the following methods of probability sampling:

i) Simple Random Sampling

This method is used when a sampling frame is small. We use random numbers to choose the sample from a listed population. First each member of the population is assigned a unique number. For example, if a population contains 1000 members, you can assign numbers 000 to 999. The random number tables are available in any standard textbook on Statistics. Choose a page, column and row at random to start selecting a sample. Suppose you want to select 30 samples out of 1000 units of a population. Since each unit of the population is identified with a 3 digit number (000 to 999). You can start selecting either the first three digits or last three digits, but not both. We have given a one page random number table at the end of this Unit. Suppose you selected fifth row and first column representing the number '998 1811', chose either '998' or '811' as first sample. In case, you have selected 811, choose the direction in which you want to move (left to right, right to left, up or down) to look for next sample unit. Suppose you have chosen from left to right select the next random number '338' from '2046338'. Continue this process till you select the required sample of 30. In case the number does not fall in the region of the population or reoccurs then simply omit it. The following random numbers are considered for 30-sample size.

811	338	574	962	229	217	192	237	399	176
165	905	523	772	710	216	880	790	358	160
955	251	800	602	190	945	009	963	604	250

The main disadvantage of simple random sampling has been observed, when the population size is very large and it is not possible to list all the members of the population.

ii) Systematic Sampling

In the systematic sampling, the sample units are selected from the population at equal intervals of time, space or order. The selection of sample in this method is very simple. Arrange all the units of the population in an order by giving serial numbers from 1 to N. Decide the sample size (n) and then sampling interval (k) by dividing the population size (N) by the sample size (n).

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

Select the first member of the sample at random from first sampling interval (1 to k). Select the subsequent sample members at equal intervals (k).

For example, if you have a population of 1000 and want to draw a sample of 100 units first arrange the population units giving serial numbers 1 to 1000. Since the sample size is 100, the sampling interval is $K = 1000/100 = 10$. Select first sample unit at random from 1 to 10. Suppose you have selected 5 then the subsequent sample units are 15, 25, 35, 45, ... 995.

The advantage of this method is that it can be used even if a formal list of population units is not available. For example, if you want to know the opinion of an ATM user, simply choose every k^{th} user (say to begin 10th) provided that we know how many users are visiting the ATM in a day (say there are 1000 users in a day and we need 100 sample size). The systematic sampling is equally precise as simple random sampling. One of the disadvantages of using systematic sampling is that if there is a periodic occurrence of the units of population then the sample selected may not be a representative sample. For example if you conduct the survey of consumers in a departmental store, the type of consumers (say housewives, working people, etc.) depends on the timing of a survey. Another disadvantage is that the first sample unit selected decides the rest of the sample units. Regardless of how we have selected the first

sample unit the rest of the sample units are automatically selected. This may lead to not giving equal chance to all population units in the sample.

iii) **Stratified Sampling**

In this method a heterogeneous population is divided into strata. Simple random sampling and systematic sampling methods are ideal if you have a homogeneous population. However, if you have heterogeneous population like male/female, rural/urban, literate/illiterate, low income/high income groups, etc. and you want to have a representative sample from each of these strata then you can choose the method of stratified random sampling. The stratified random sampling is in fact a simple random sampling of each stratum of the population. For example, in a study of different income groups, a simple random sample may be drawn for each income group (either in proportion or disproportion to the stratum population size). The main advantage of stratified random sampling is that since the random sample is drawn from each stratum of the population, this method is more representative and thus represents more accurately the characteristics of the population. The drawback of the stratified sampling is that it is often difficult to prepare a list of population members for each stratum since the lists may not be readily available.

iv) **Cluster Sampling**

In cluster sampling we divide the population into smaller groups called clusters and then select samples randomly from one or more groups (clusters). Here we assume each cluster of population will have all the characteristics of the population as a whole. The cluster sampling is more appropriate in geographical studies. For example, suppose you are conducting a survey in Mumbai. Instead of spreading the sample to the whole geographical area of Mumbai, you can divide the city into say 30 blocks, select a few blocks (say 3 or 4) and then select a representative sample from each block. Selection of groups or clusters of elements rather than individual elements is regarded as a more practical approach to surveys it samples. However, the accuracy of the survey may decline if the clusters are not representative of the population characteristics. Sometimes, the cluster sampling is also called as two-stage sampling.

v) **Multi-stage Sampling**

Multi-stage sampling is a generalisation of the cluster sampling. In cluster sampling what you have observed is a two-stage process. In the first stage divide the population into a number of clusters, and in the second stage from each of the sampled clusters you choose a representative sample. Multi-stage sampling is, thus, carried out in a number of stages. Progressively smaller population is selected randomly at each stage. Further, at each stage of the multi-stage sampling you can use a suitable method of sampling (say stratified, simple random, or systematic, or cluster). For example, if you are conducting a survey covering all the geographical area of India. You can choose a representative sample of states in the first stage, and in the second stage you can select a representative sample of districts from each state, and so on. The main advantage of multi-stage sampling is that it reduces the costs of data collection when the survey is in-person. The drawback of the multi-stage sampling is that if the sampling units selected at different stages are not representative, the findings of the survey have a lower value in terms of their generalisability.

vi) **Non-Probability Sampling Methods**

The non-probability sampling methods are widely used as a case selection method in qualitative research or for preliminary and exploratory nature of quantitative studies where random sampling is too expensive. However, the probability methods of sampling are always preferable and non-probability methods of sampling are not a substitute to the former. In this section,

we cover three non-probability sampling methods, namely, convenience sampling, judgment sampling, and quota sampling.

- **Convenience Sampling:** In this sampling, the researcher selects the samples that are convenient or on the basis of availability. For example, in conducting opinion polls people who are coming out of polling booth and are willing to talk are chosen for interview. Mail-in surveys in newspapers, magazines or TV are other examples of convenience sampling.
- **Judgment Sampling:** In judgment sampling, the researcher interviews a panel of individuals known to experts in a field. This sampling method is also sometimes called purposive sampling. The sample is selected entirely on the basis of researcher's/expert's judgment on the characteristics of the sample. The consumer price index is based on a judgment sample. In the absence of representative data or where it is difficult to establish the universe for a sample one may resort to judgment sampling.
- **Quota Sampling:** Quota sampling is similar to convenience sampling, except that the proportionality of each stratum is preserved. The samples are selected on the basis of some parameters like sex, age, region, income, etc., in order to make a representative sample. The researcher then chooses a fixed quota from each of the sub-groups. Unlike stratified random sampling, in quota sampling the samples are not drawn randomly but a fixed quota is chosen. For example, a market survey may cover 5 market centres of a city and from each market centre a fixed quota of 50 consumers can be interviewed. Their selection is not done randomly, but only according to the researcher's convenience.

21.2.6 Sample Size Determination

How large the sample should be is a question that arises at the planning stage of any survey. If you take a larger sample than required then it results in wastage of resources because you will not gain any additional information. On the other hand, if you have too small a sample than required, you may not get reliable results, and hence they are of no practical use in making a good decision. The main objective in sampling is to obtain a desirable accuracy and desirable confidence level with minimum cost. All probability-sampling methods ensure representative sampling. Also, technically speaking, a sample size is related to the confidence level or precision level. The typical confidence levels are 95 per cent, and 99 per cent and typical precision levels are 5 per cent, and 1 per cent. The discussion on how to estimate the sample size is not covered here since it is beyond the scope of this Unit.

21.3 COMPUTER SIMULATION

Simulation is a method of solving problems in decision-making by developing a mathematical model of a process and then conducting a series of experiments to predict the behaviour of that process. In other words, simulation is used to describe a procedure of establishing a model and deriving solution numerically. We use a simulation model to conduct experiments.

A mathematical model can be used either for the purpose of optimisation or description. Using optimisation models you can seek points of maximisation or minimisation. For example, a policy analyst often would like to know what conditions or actions would result in profit/revenue maximisation or cost minimisation. On the other hand, a description model describes the behaviour of a system. It only indicates a valid representation of reality and does not directly suggest optimising conditions. From this, you can understand under what conditions a system performs more efficiently. In simulation, you use a model for optimizing purposes by raising 'what if' questions. For example, as a policy analyst you would like to know 'what will be the effect on the benefits (in terms of

money) if the cost of a project increases by 5 per cent or 10 per cent or 15 per cent and so on? Using simulation you can obtain useful information for optimisation of the project.

Simulation models are typically used for the analysis of complex problems. We will explain this with a simple example. Suppose you are asked to find the sum of the numbers 1 to 20

(1+2+3+4...+20). Using Arithmetic Progression formula $\frac{n(n+1)}{2}$ you find the sum of n terms. In

our example, you have n=20 therefore sum of 1 to 20 is $\frac{20(20+1)}{2} = 210$. In case you are not

familiar with the formula, what do you do? You may perhaps start adding 1+2=3 and then add 3 to the resulting value (3) and continue the process till you complete adding all the 1 to 20 numbers. Suppose you are asked to add 1 to 100. In this case, imagine how cumbersome it is and how much time it takes if you follow the second method. In fact, the first method of using formula is more efficient. However, there are complex situations where you may not use direct formulas, but try to achieve them through experiments.

21.3.1 A Simple Simulation Model

We will illustrate the concept of simulation with a simple example. Suppose a distance education institution is offering a postgraduate programme to its students. In relation to the number of students enrolled, the institution incurs Rs.60 millions as a fixed cost on developing course materials and other overheads. It also incurs Rs.500 per student as variable cost for every additional enrolment. In order to find the total cost for different enrolments you may use the following formula.

$$TC = FC + VC \times S$$

Where TC = Total Cost

FC = Fixed Cost

VC = Variable Cost per student

S = Number of students

Now suppose you are given FC = Rs.60 millions and VC = Rs.500 per student, you would like to know what will be the total cost for different levels of enrolment. Using the above forecasting formula you may project the total costs for different levels of enrolment as shown in the following table.

Table: 21.1: A Total Cost Simulation

Number of Students	Fixed Cost (Rs.)	Variable Cost Per Student (Rs.)	Total Cost (Rs. in lakhs)	Average Cost Per Student (Rs.)
1000	600,00,000	500	605,00,000	60500
5000	600,00,000	500	625,00,000	12500
10000	600,00,000	500	650,00,000	6500
15000	600,00,000	500	675,00,000	4500
20000	600,00,000	500	700,00,000	3500
25000	600,00,000	500	725,00,000	2900
30000	600,00,000	500	750,00,000	2500

Further, you may be interested to find out what will be the average cost per student if the following formula is applied.

$$AC = \frac{TC}{S}$$

Where AC = Average Cost

With these simulations as a policy analyst you may understand that with marginal costs per student you are able to achieve greater economies of scale for additional enrolments. You may conclude that higher level of enrolments in distance education will result in higher level of cost efficiency.

21.3.2 Reasons for Using Simulation

Simulation is one of the most frequently used techniques for management decisions. Linear programming problems, input-output analysis, cost-benefit analysis, and forecasting techniques are some of the examples of using simulation. While conducting cost-benefit analysis or solving a linear programming problem, the policy analyst may try to explore different policy alternatives incorporating different perceptions and attitudes about risk and uncertainty. Here, the policy analyst's goal is to clarify the issues and put forward the consequences of a variety of alternatives. This will help the decision makers to choose one of the alternatives that match their values and objectives.

Generally, simulation is used in conditions where mathematical formulation and solution by the use of models are not feasible. Simulation has been used in numerous decision-making problems, ranging from inventory and queuing management to energy policy modelling. Also, simulations are used to analyse large and complex real world situations that cannot be solved by conventional mathematical models. Another advantage of simulation is that it is possible to experiment on a system without exposing the organisation to real world dangers. By exploring the changes in a real world system through simulation you can learn how to improve the system without trying both bad and good proposals on hand.

The decision maker's past experience and intuition are significant factors for successful decisions. Often they have to manage complex systems. Simulation aids the policy decision makers past experience and intuition that ultimately may result in a successful decision.

21.3.3 Limitations of Simulation

In spite of several advantages of simulation, there are some limitations in using simulation for decision making. Some of the drawbacks are:

- Simulation by itself does not provide the answer to a problem. It only provides a technique for evaluating possible answers.
- Simulations do not generate optimal solutions like other mathematical models. It is only a trial and error approach that may produce different solutions.
- The simulation for probabilistic systems is susceptible to sampling error.
- The development of a large simulation model can be expensive and can take longer time. For example, it takes years to develop a good corporate planning model.

21.3.4 Computer Simulation

With the advent of computers, simulation models are widely used to gain information useful for decision-making. Computer spreadsheet software like Microsoft Excel and Lotus 1-2-3 are used, for financial simulations. The power of these spreadsheets is to do "what if" analysis. The spreadsheets also have a random number generator that supports probabilistic simulations. We will give (section

5. Formulae are entered to perform a range of arithmetic calculations. For example, =SUM(A4+B4) means the sum of two numbers located in cells A4 and B4.

The power of these spreadsheets is to do 'what-if' analysis. For example, suppose you are involved in the preparation of a budget of a state government. In the process you are interested to know how much revenue can be generated through sales tax on petrol. Suppose, the current sales tax is 30 per cent of the cost of petrol and it is likely that there will not be any change in the sales tax in the coming budget. There are two ways you can expect additional revenue through this tax. These are if and only if either cost of petrol or unit sales of petrol or both increases. However, you may not exactly know how much these two units increase and want to experiment with different rates. This can be done using the spreadsheet and conducting 'what-if' analysis. This is presented in the following figure.

The screenshot shows a Microsoft Excel spreadsheet titled 'Book11'. The spreadsheet contains a 'what-if' analysis for petrol sales and costs. The data is organized as follows:

	Current	2%	5%	7.5%	10%
1 Assumptions					
2 Petrol sales increases by	Current	2%	5%	7.5%	10%
3 Petrol cost increases by	Current	5%	10%	15%	20%
4 Current year unit sales (in litres) =	100,000,000	102,000,000	105,000,000	107,500,000	110,000,000
5 Current unit cost (Rs./litre) =	25.00	26.25	27.50	28.75	30.00
6					
7 Petrol sales including sales-tax (Rs.)	3,250,000,000	3,480,750,000	3,753,750,000	4,017,812,500	4,290,000,000
8 Petrol sales excluding sales-tax (Rs.)	2,500,000,000	2,677,500,000	2,887,500,000	3,090,625,000	3,300,000,000
9 Revenue (Rs.)	750,000,000	803,250,000	866,250,000	927,187,500	990,000,000

The likely scenario of revenue through sales tax is also presented in the above Figure. You may also be interested to see a different scenario like absence of increase in the cost of petrol or sales volume or a different rate of sales (or cost) increase (or decrease). This will take less than a minute to see the new estimates. It is very easy to perform 'what-if' analysis using spreadsheets.

Simulation Using a Special Purpose Language

A special purpose simulation language called SIMSCRIPT was developed by the New York City-Rand Institute (Grace Carter, 1974) to optimise the development of fire fighting units. The decision variables here include how many fire-fighting units are required to operate, where to deploy them, and how to dispatch them. The SIMSCRIPT language uses three programmes, namely, (a) incident generator, (b) simulator, and (c) compares. The 'incident generator' creates a data file on the sequence of alarms. For each alarm the data continues when and where the alarm occurs, what type of fire fighting equipment is required, and how long the equipment is required. The 'simulator' is the model defined by the 'incident generator'. The model captures the information on dispatch of vehicles, servicing, and returning to the fire station to await the next call. Finally, the third programme 'compares' the information generated by the simulator alarm-by-alarm basis on the assumption of number of fire stations, their locations, and dispatch rules.

The above model was used by David and Thomas (1977) to analyse the fire suppression system of Denver, Colorado. The simulation analysis performed recommended a reduction of the number of fire fighting stations from 44 to 39. This includes closing down obsolete stations and building

better ones at new locations. The recommendations resulted in annual cost avoidance of \$1.2 millions and accumulated net cost reduction of about \$2.3 millions over a 7-year period.

21.4 CONCLUSION

This Unit introduced a number of methods for conducting surveys used frequently in policy analysis. You have learned the differences between cross-sectional surveys and longitudinal surveys. The Unit attempted to answer questions, such as, What is a sample and population? Why samples should be drawn in such a way that they represent all the characteristics of the population? What are the considerations in deciding a sample size?

Four methods of survey are discussed. These are: In-person interview, Telephone survey, Mail survey, and Internet survey. Each method has its own advantages and disadvantages. After going through this Unit, you should be now confident of choosing any of the methods suitable for your Survey. These days the mailed survey and Internet survey are gaining much importance because of their cost-effectiveness, time saving, and advancements in the application of information technology.

The questionnaire-design is another important aspect of any survey. Some guidelines are provided in designing questionnaires that will be helpful in improving the response rates.

The sampling methods can be classified into: probability sampling and non-probability sampling methods. The probability sampling methods are based on the equal chance of each population unit, getting included in a sample. The non-probability methods are not in support of representative samples, but they are often used because of their convenience and there is no other way of selecting a sample for conducting a survey.

In this Unit, you have learned what is simulation, and why do you need simulation in policy decisions. In addition an analysis of advantages and drawbacks of simulation is also provided. These will help you to understand before actually using simulation. Two typical examples are given to provide a conceptual framework for simulation models. You can either use the general-purpose computer languages like FORTRAN, Pascal, C/C++, etc. or special purpose simulation languages like GPSS, SLAM, GEMS, SIMSCRIPT, etc. However, each one of them has its own advantages and disadvantages.

21.5. KEY CONCEPTS

- | | |
|----------------------------|--|
| Cluster Sampling | : We divide the population into groups called clusters and then select sample units randomly from one or more clusters. This method is also called two-stage sampling. |
| Cohort Studies | : Cohort studies focus on the same population over a period of time, but the difference is that the samples are drawn from the same group, which has experienced an event or participated in a programme. |
| Convenient Sampling | : In this method of non-probability sampling, the sample units are selected according to the convenience of the researcher/ interviewer. |
| Judgment Sampling | : In this non-probability sampling method the samples are selected according to an expert's/researcher's judgment, which perceives that the sample units have all the characteristics of the population of the universe. |

Longitudinal Surveys	: Longitudinal surveys are used to collect data on a population or sub-groups over a period of time.
Multi-stage Sampling	: This probability sampling method is an extension of cluster sampling where there are more than two stages.
Questionnaire	: Apse-designed format of questions used for collecting data in survey research.
Quota Sampling	: In this non-probability sampling method the sample units are selected according to some quotas after grouping of the population units based on characteristics, such as, age, sex, region, income, education, etc.
Random Sample	: An unbiased sample that consists of all the characteristics of a population of the universe.
Sample	: A sample that represents all the characteristics of the population of a universe.
Simple Random Sampling	: A probability sampling method where each unit of population has equal chance of inclusion in the sample.
Simulation	: Simulation is a method of solving decision-making problems by developing mathematical models and then conducting a series of experiments to predict the behaviour of interacting variables.
Stratified Sampling	: In this sampling method, the heterogeneous population is stratified to make homogeneous groups (stratums) for the purpose of sampling.
Systematic Sampling	: In this probability sampling method the sample units are selected at equal intervals in terms of time, space, or order provided that the first sample unit is selected at random from the first k (interval) units of population.

21.6 REFERENCES AND FURTHER READING

- American Statistical Association, 1997, "Survey Research Methods", at www.stat.ncsu.edu/info/srms/srms/html <<http://www.stat.ncsu.edu/info/srms/srms/html>>.
- Babbie, Earl R, 1973, *Survey Research Methods*, Wadsworth Pub. Co, Belmont, CA.
- Fowler Floyd, J. Jr., 1988, *Survey Research Methods*, C A Sage Publication, Newbury Park.
- Goode and Hatt, 1952, *Method in Social Research*, International Student Edition, Mc Graw Hill.
- Gordon, Geoffrey, 1987, *System Simulation*, Prentice-Hall of India, New Delhi.
- Grace, Carter, 1974, *Simulation Model of Fire Department Operation: Program Description*, The New York City-Rand Institute, New York.
- Hugh, J. Watson, and John H. Blackstone Js. 1989, *Computer Simulation*, John Wiley & sons, New York.
- Kerlinger, F.N., 1986, *Foundations of Behavioural Research*, TX: Holt, Rinehart and Winston, International Edition, Fort Worth.

Kothari, C.R., 1989, *Research Methodology: Method and Techniques*, Wiley Eastern Limited, New Delhi.

Monarchi, David E., Thomas E. Hendrick, and Donald R. Plane, "Simulation for Fire Department Deployment Policy Analysis", *Decision Sciences*, January, 1977.

Sadhu, A.N. and A. Singh, 1980, *Research Methodology in Social Sciences*, Sterling Publishers Private Limited, New Delhi.

Shenoy, G.V.S., et.al, 1983, *Quantitative Techniques for Managerial Decision Making*, Wiley Eastern, New York.

21.7 ACTIVITIES

- 1) When are simulation models more appropriate than optimisation models? Discuss.
- 2) Discuss the advantages and disadvantages of general-purpose computer languages and special purpose simulation languages for conducting simulations.
- 3) Consider the following total revenue function:
 $R = 45 - 0.5Q^2$ where R is the total revenue and Q is the Quantity sold.
 - a) Using optimisation concept (discussed in Unit 22) find out the value of Q that maximises R .
 - b) Suppose you do not know the differential calculus, how do you solve the above problem using a computer? Explain the mode of analysis performed.
- 4) Discuss the differences between trend studies, cohort studies, and panel studies with suitable examples of policy analysis. What are the limitations of panel studies?
- 5) Discuss the merits and drawbacks of various methods of survey. Why are mailed surveys and Internet surveys gaining importance these days?
- 6) Why are probability-sampling methods preferred to non-probability sampling methods? Discuss.
- 7) What are the issues that you consider while designing a questionnaire? Is it important to have a larger sample size? Discuss.

Random Number Table

Column

	1	2	3	4	5	6	7	8	9	10	
Row	1	8321960	3738691	9799403	2048969	5040987	9644530	1043990	5140028	2331956	425679
	2	5455861	8056250	2544572	2352713	8108579	6463133	6545800	8418443	9640450	097751
	3	4124798	2705391	1761560	2390302	7271422	4745529	7958358	1142812	5671381	074349
	4	1756223	4000127	7361071	5190949	2421704	5666414	1964120	1736970	0898522	650136
	5	9981811	2046338	9202574	0588962	6306229	5587217	4168192	7922237	6404399	971111
	6	3483165	5570905	3969523	6801772	2319710	3982216	1458880	9772790	0686358	283416
	7	0273955	6402251	8623800	6450602	3477190	1591945	6604009	0084963	6535604	827621
	8	2623257	3169089	5734610	1713259	8034166	5654825	8150190	5130311	6897834	574450
	9	5116188	8891882	9287974	9306457	0824236	6936207	2603345	9568262	2010966	180959
	10	8192202	4158078	8597397	7153642	8035351	7294605	2431042	7375233	4280501	091388
	11	7728407	7358939	4251412	5232607	1542070	6781856	6723336	2000772	1691755	684258
	12	2796466	3168988	4740748	1003690	9357163	8809737	2540113	8564549	4549039	822227
	13	6202090	3028511	5119455	0283291	4316705	6468604	0270553	9215994	2388410	869018
	14	2842317	8369317	7179585	2817675	4501995	6208468	5479155	2895887	2616006	391248
	15	6509176	6647923	1361518	3585571	6002538	3034686	5759994	1011945	0428550	089987
	16	8728736	9827551	8489409	6601482	5270991	8514979	7722147	1321316	8755556	043107
	17	7027408	5705233	3206106	9276893	0474701	4147247	2230458	7446013	9409230	016001
	18	8289615	3571820	0957198	2756292	1656700	2560460	8392388	3760152	0568193	233415
	19	4420793	5882651	5627136	6216428	3649258	1473337	1755096	8908170	9254025	589354
	20	8725326	2591197	0263556	0499703	9692758	4684871	2227638	9745705	4686535	856926
	21	3543878	0870871	4937410	2417481	4193760	9027624	2407109	2328001	5411904	241870
	22	6542864	6236422	0011015	0295963	8769055	2997607	1611662	0250323	0538884	449352
	23	9583781	3045137	8407762	4743086	3956983	3639166	2360412	3721797	3881206	943997
	24	4425626	2091448	2223679	0867541	2228645	9636460	1789379	5600538	4994108	924485
	25	6184612	5847113	6608373	2943696	2155685	3094922	7340263	0838375	5794133	629392
	26	1191788	4731439	9579476	2147703	4707867	8070442	1919593	4432746	6112606	401582
	27	6667402	1538925	5897065	7186643	2130158	2399644	9013726	5915962	2774931	470973
	28	2669584	6426328	5079538	8251857	3599028	8171955	8882381	5973371	6585703	030608
	29	0093668	8142995	4465615	8916068	6818559	7183686	3093475	3134759	8581724	322765
	30	2026517	0922442	1412703	6266272	3175006	3842448	8796268	0262526	0381968	342588
	31	4534024	7697053	4868217	2868737	5974617	8823016	2949217	6369860	0505759	917430
	32	0760738	2913215	2793140	3540176	0177805	8334427	4521757	0287290	3591531	453879
	33	3325134	7633260	8364478	9656036	4047642	4009275	7199712	0483862	4124769	147271
	34	8256482	5655913	8839461	4459925	3040043	8377818	0661094	4715872	6695119	488058
	35	9025690	4298285	5574813	2354160	2322047	6938196	2137818	7625460	1752005	855690
	36	3419285	9438115	3838089	5986563	6148386	5410764	1225104	3858088	2391297	704101
	37	2912915	1927810	5139550	7635561	0957373	4470026	2239259	3873148	0294959	890095
	38	9851067	8343236	1033181	8982930	3467555	6966724	2312696	5103618	4742406	727836
	39	9440894	2943427	9510627	5980986	1380771	2665220	7242862	4263393	2892430	867053
	40	1137867	5437166	523544	3842000	4398343	8240530	2355967	66642	6605037	18543
	41	7109003	3313227	1314113	8019056	7260747	9904799	7139823	2318596	3042956	698504
	42	9630926	2208324	66570	5900743	3305479	5226677	3312048	9735251	2013291	464281
	43	7439586	9182637	4795841	2068817	9962328	2133122	1501738	678718	7806841	310339
	44	6284260	538363	9377852	2002132	5048511	3541036	8629627	3682464	7590586	926320
	45	9834361	6063463	1380169	6755535	8223905	9081534	1862991	8544677	9388239	400809
	46	5371604	5618530	1701089	8507762	6842081	8187969	6154011	655217	6998747	680708
	47	6633242	4480982	9836063	2084383	2841356	7070790	4337341	9142806	3614065	584243
	48	7182588	692296	1147474	1451427	6581653	530472	1876589	2098062	7183538	163695
	49	2310988	8626979	7669620	5454810	3800372	5567410	9650555	211692	1137645	492244
	50	6774333	2501281	7189333	8555681	7452621	4303975	1446997	3065976	3067915	210209