Indira Gandhi
National Open University
School of Social Sciences

# MEC-109
# Research Methods in Economics

Block

# 2

## RESEARCH DESIGN AND MEASUREMENT

## Expert Committee

Prof. D.N. Reddy
Rtd. Professor of Economics
University of Hyderabad, Hyderabad

Prof. Romar Korea
Professor of Economics
University of Mumbai, Mumbai

Prof. Harishankar Asthana
Professor of Psychology
Banaras Hindu University
Varanasi

Dr. Manish Gupta
Sr. Economist
National Institute of Public Finance and Policy
New Delhi

Prof. Chandan Mukherjee
Professor and Dean
School of Development Studies
Ambedkar University, New Delhi

Prof. Anjila Gupta
Professor of Economics
IGNOU, New Delhi

Prof. V.R. Panchmukhi
Rtd. Professor of Economics
Bombay University and Former
Chairman ICSSR, New Delhi

Prof. Narayan Prasad (**Convenor**)
Professor of Economics
IGNOU, New Delhi

Prof. Achal Kumar Gaur
Professor of Economics
Faculty of Social Sciences
Banaras Hindu University, Varanasi

Prof. K. Barik
Professor of Economics
IGNOU, New Delhi

Prof. P.K. Chaubey
Professor, Indian Institute of
Public Administration, New Delhi

Dr. B.S. Prakash
Associate Professor in Economics
IGNOU, New Delhi

Shri S.S. Suryanarayana
Rtd. Joint Advisor
Planning Commission, New Delhi

Shri Saugato Sen
Associate Professor in Economics
IGNOU, New Delhi

## Course Coordinator and Editor: Prof. Narayan Prasad

## Block Preparation Team

| Unit | Resource Person | Editor (Format, Language and Content) |
|---|---|---|
| 6 | Prof. Narayan Prasad<br>Professor of Economics<br>IGNOU, New Delhi | Prof. Narayan Prasad<br>Professor of Economics<br>IGNOU, New Delhi |
| 7 | Sai S.S. Suryanarayan<br>Rtd. Joint Advisor<br>Planning Commission, New Delhi | Prof. Narayan Prasad<br>Professor of Economics<br>IGNOU, New Delhi |
| 8 | Prof. A.K. Gaur<br>Professor of Economics<br>BHU, Varanasi<br>&<br>Prof. Narayan Prasad<br>Professor of Economics, IGNOU, New Delhi | Prof. H.S. Asthana<br>Professor of Psychology<br>BHU, Varanasi |

## Print Production

Mr. Manjit Singh
Section Officer (Pub.)
SOSS, IGNOU, New Delhi

October, 2015



ISBN-978-81-266-



Printed and published on behalf of the Indira Gandhi National Open University, New Delhi by the Director, School of Social Sciences.

Laser Typeset by : Tessa Media & Computers, C-206, A.F.E.-II, Okhla, New Delhi
Printed at :

# BLOCK 2 RESEARCH DESIGN AND MEASUREMENT

Research design is a logical structure of an enquiry and its formulation is guided and determined by the research questions raised in the problem under investigation. Apart from specifying the logical structure of data, research design also test and eliminate alternative explanation. Broadly, the observational design sampling design and statistical design are covered in Research Design. The various attributes of people, objects or concepts are being increasingly included in explanation of human behaviour in Economics. Hence, these individual traits, attitudes need to measure for deeper analysis. Research Design and measurement issues, therefore constitute the theme of this block. The block comprises of 3 Units.

Throwing light on the concept of research design and various types of research, **Unit 6** covers the rationale and forms of mixed methods research with illustration of three case studies.

**Unit 7** provides comprehensive knowledge of all the elements for collecting the quantitative data for research study. It covers the methods and tools of data collection and the various issues related to sampling design.

**Unit 8** deals with the various measurement scales, scaling techniques, and criteria for good measurement and sources of errors in measurement.

# UNIT 6 RESEARCH DESIGN AND MIXED METHODS RESEARCH

**Structure**



## 6.0 OBJECTIVES

After going through this Unit, you will be able to:

- state the various types of research;
- explain the concept and meaning of research design;
- make a distinction between research design and research methods;
- delineate the characteristics of quantitative research and qualitative research;
- discuss the purposes of mixed methods research; and
- appreciate the application of mixed methods research in conducting research studies in social sciences in general and economics in particular.

## 6.1 INTRODUCTION

Due to confusion or lack of clarity between research design and research methods, we often use these two terms interchangeably. As a result, the research design is evaluated on the basis of strengths and weakness of the method rather its ability to draw relatively unambiguous conclusions. Hence

efforts have been made here to throw light on the distinction between these two terms. Traditionally two approaches – quantitative research (QN) associated with post positivist paradigm and qualitative research (QL) associated with interpretative paradigm have been followed in social sciences. However, in recent years increasing reliance on mixed methods research by combining quantitative and qualitative methods in a single research project has been observed. Hence in this unit, the concept of mixed method research, its types and application have been covered. Application of mixed method research in a subject like economics which is essentially quantitative in nature have been illustrated by case studies. Before taking up the issue of research design and research method, it is desirable to have an idea about various types of research. Hence let us begin to introduce the various types of research.

## 6.2 TYPES OF RESEARCH

There are different 'types of research' depending on their nature and field of specialization. While a classification based on a dichotomous distinction (like, theoretical/applied, descriptive/analytical, conceptual/empirical, etc.) is possible, it is necessary to recognize that there may be overlapping features rendering such a classification less perfect to that extent. It is, therefore, useful to first of all take a look at the underlying basic features so as to be able to identify the different types of research within their respective connotations and usage.

A variable which can be measured can take values or scores either in quantitative or qualitative terms. For instance, yield of an agricultural output, height/weight of individuals, etc. can be measured and expressed in quantitative terms. On the other hand, characteristics like one's feelings or opinion (e.g., good/bad, male/female, agreeing/disagreeing, yes/no, etc.) are attributes on which a choice can be expressed (and a score assigned) depending on the possible alternatives or choices. Thus, variables like the performance of an artist, respondent's gender, etc. on which responses can only be categorized or grouped are **'qualitative variables'**. From this angle, variables on which data can be collected and expressed in quantified terms may be called as '**quantitative variables**'. Even variables which are expressed in qualitative terms can be counted whereupon their total numbers become quantified expressions. A popular example of developmental economics which ranks different countries based on their status attained in a number of areas/variables(some of which like the political freedom/choice enjoyed by the country, though expressed in some quantitative measure, is qualitative in its nature) is the human development index. In actual research (particularly research in the field of social sciences), it is important to note that variables of both the nature normally co-exist. In view of this, the analysis pertaining to each type of variable/data needs different methods and techniques.

The type of research would also vary depending on the objectives of the study. Research design varies with the type of research one likes to pursue. With this background, we can now proceed to know the distinctions between different types of research.

### 6.2.1 Theoretical and Applied Research

Research can either be theoretical or applied. The former i.e., theoretical research can also be considered as 'fundamental research' as its outcome serves

as a foundation for all subsequent development in the field. Fundamental (or basic) research mainly concerns with formulation of theory with knowledge perceived as an end in itself. It, thus, aims at obtaining knowledge of the logical processes involved in a phenomenon. It pertains to the quest for knowledge about a phenomenon without concern for its practical use. Such a research may either verify the old theory or establish a new one. For example, fundamental research in economics may consist of research to develop and improve economic theories or evolve quantitative techniques to measure parameters such as multiplier effect, elasticity of demand and supply, etc. Fundamental research is essentially positive in nature.

Applied research, on the other hand, aims at finding a solution for a problem facing society or industry. It is, thus, applied to practical situations or contexts. While pure research discovers principles and laws, applied research discovers ways of applying them to solve specific problems. It is useful to test the theories developed empirically and can as a result also contribute to improving the tools and techniques of measurement. Basic research, therefore, can be treated as building blocks for applied research. Illustrations of applied research in economics can be measurement of poverty, employment, rural development, agriculture, environment, etc.

### 6.2.2 Descriptive and Explanatory Research

Descriptive research describes a situation, events or social systems. It aims to describe the state of affairs as it exists. Surveys and fact-finding enquiries of different kinds are part of descriptive research. Survey methods of all kinds including comparative and correlational methods are used in descriptive research studies. A survey of socio-economic conditions of rural/urban labour is an area of descriptive research. In descriptive research studies, the researchers have no control over variable. They can report only what has happened or is happening. Descriptive research deals with questions like 'how does X vary with Y? or 'how does malnutrition vary with age and sex?, etc.

Explanatory research, aims at establishing the cause and effect relationship. The researcher uses the facts or information already available to analyse and make a critical evaluation of the data/information. An example of explanatory research is: 'whether increase in agricultural productivity is explained by improved rural roads?'

### 6.2.3 Quantitative and Qualitative Research

**Quantitative Research** is the systematic and scientific investigation of quantitative properties and phenomena and their relationships. The objective of quantitative research is to develop mathematical models, and to test hypotheses. Based on ontological and epistomological premises, quantitative research is characterized by the following attributes:

- A belief in a single reality.
- The pursuit of identifying universal laws.
- Separation of knower (researcher) from known (researched).
- The possibility and necessacity of value free research.
- Pursuit of universal laws (findings) beyond the limit of research/social context.

- The tendency to work with large, representative sample.
- An emphasis on deductive research via falsifiable hypothesis.
- Formal hypothesis testing.

In contrast, **Qualitative Research** captures a set of purposes associated with meaning and interpretation. Stress is laid on the socially constructed nature of reality, the intimate relationship between the research and researched and situational constraints that shape the enquiry. The key attributes associated with qualitative research are:

- A belief in a constructed reality, multiple realities, co-existence realities.
- An interdependence between the knower and the known.
- Impossibility to separate the researcher from the research subject.
- Heavy role of context in research process.
- The impossibility to generalize research findings beyond the limits of the immediate context.
- Non-separation of cause and effect.
- The explicit focus on inductive, exploratory research approaches.
- The tendency to work with small and purposely chosen sampling.
- Analyses holistic system.

### 6.2.4 Conceptual and Empirical Research

Conceptual research is related to abstract ideas or theory. It is generally used by philosophers and thinkers to develop new concepts or to reinterpret the existing theories.

Empirical research relies on experience or observation. It is data based. It is subject to verification by observation and experiment. This type of research is particularly useful when validation or verification of an aspect is required.

### 6.2.5 Other Types of Research

Research may be exploratory or formalized. **Exploratory** research aims at developing the hypothesis rather than testing a pre-conceived hypothetical contention or notion. **Formalised** research studies deal with a definitive structure within which specific hypotheses are tested. **Historical research** utilizes existing documents to study events of the past. Research can also be experimental or evaluative. **Experimental research** aims at identifying the causal factors by means of experiments. In **evaluative research**, the cost effectiveness of a programme is examined. Such research addresses the question of the efficiency of a programme and are useful in taking policy decisions on issues like whether the programme is effective and/or needs modification or re-orientation? Whether it should be continued?

**Action research** is another type of research. It deals with real world problems aimed at finding out practical solutions or answers to them. It gathers feedback which is then used to generate ideas for improvement. You will find details on Action Research in Unit 20.

**Check Your Progress 1**

1) Distinguish between quantitative variable and qualitative variable.

……..

……..

……..

2) State the main objectives of applied research?

……..

……..

……..

3) In what sense quantitative research is different from qualitative research?

……..

……..

……..

4) What is the distinction between explanatory and descriptive research?

……..

……..

……..

## 6.3 RESEARCH DESIGN

Research design is a logical structure of an enquiry. Given the research question or theory, what type of evidence is needed to answer the question (or to test the theory) in a convincing way – constitutes the essence of the research design. Let us use an analogy to understand the term 'research design'. While constructing a building, the first decision to be arrived at is: whether we need a high rise office building, a factory, a school or a residential apartment etc.? Until this is decided, we cannot sketch a plan and order material or setting critical dates for completion of the project dates. Similarly, a social researcher needs to be clear about the research questions and then the research design will flow from the research questions. The function of a research design is to ensure that the evidence obtained enables us to answer the initial research questions as unambiguously as possible. Obtaining relevant evidence entails specifying the type of evidence we need to answer the research question, to test a theory, to evaluate a programme or to accurately describe some phenomenon. The issues of sampling, method of data collection (e.g. questionnaire, observation, document analysis), design of questionnaire etc. are all subsidiary to what constitute the evidences that need to collect to answer the research questions.

Thus the research design 'deals with a logical problem and not a logistical problem. Apart from specifying the logical structure of the data, it also aims to test and eliminate alternative explanation of results.

The research design comprises of the following components:

i) the sampling design; (the type of sampling method i.e. random or non random sampling);

ii) the statistical design (i.e. the sample size and the method to draw the sampling to be adopted);

iii) the observational design (i.e. the instrument of collection of data);

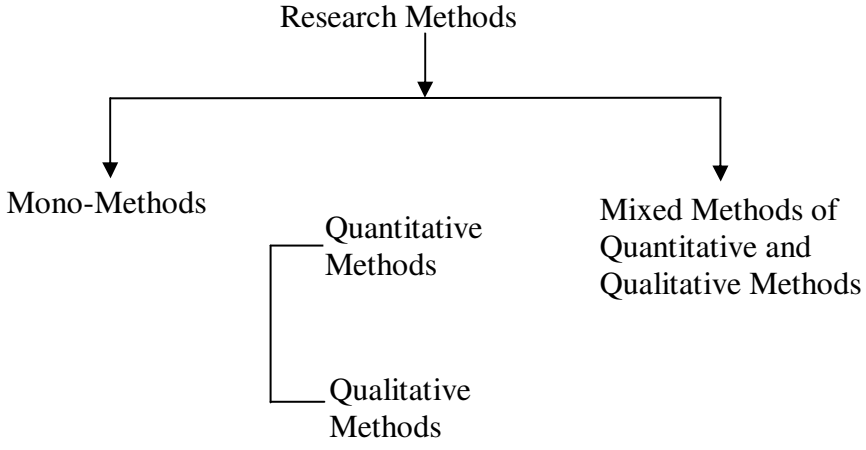
iv) the operational design i.e. the specific details by which the procedure in (i), (ii) & (iii) above are to be carried out.

## 6.4 RESEARCH DESIGN VS. RESEARCH METHODS

Research design is different from the research methods. The research methods are made of data collection and techniques of data analysis whereas the research design is a logical structure of an enquiry. There is nothing intransic about any research design that requires a particular method of data collection. How the data are collected is irrelevant to the logic of the research design. Confusing research designs with research methods leads to poor evaluation of designs. For example equating cross sectional designs with questionnaires or case studies with participant observation result in evaluation of research designs against the strengths and weakness of the method rather than their ability to draw relatively unambiguous conclusion or to select between rival plausible hypothesis.

## 6.5 RESEARCH METHODS

Research methods as explained in Unit 1 refers to tools and techniques of data collection and data analysis. Broadly Research Methods are put under two categories: (i) Mono Methods, (ii) Mixed Methods. Within mono methods, there are two approaches (i) quantitative (QN) methods (ii) qualitative (QL) methods. Since quantitative methods are associated with positivist and post-positivist paradigm, quantitative method is also termed as (QN) method research. Similarly by virtue of the fact that QL methods are associated with interpretativist paradigm and critical theory paradigm, research conducted by the qualitative method is also termed as qualitative methods research.



### 6.5.1 Quantitative Methods

The researcher using this approach tests theory through empirical observations and inclines to establish cause and effect relationship. Well predesigned structured questionnaire and structured interviews are resorted as tool to collect the data. The variables that can be measured are emphasized. The descriptive, analytical and inferential statistical techniques are used to analyse the data. Experimental research design is used to make the logical structure of the research study. Generalization of laws and their replication is emphasized.

### 6.5.2 Qualitative Methods

Under qualitative approach, a researcher starts with observation as a basis for generating theory and concentrates on meaning of observations. He studies events as they occur in naturalistic settings. He allows interview topics to emerge during conversation and listens to others' interpretation and perspectives.

Under qualitative method, open-end questionnaire, key informants, group discussion and unstructured interviews, documents, interview transcripts etc. are used as tools to collect data. Content analysis, Case study, Action research participatory method, Cluster analysis, Factor analysis, Correspondence analysis, Context analysis, are used in analyzing qualitative data. You will find details of these techniques in Unit 13, 14, 15, 16 and 17 of Block 4 and 18, 19 and 20 of Block 5 of this course.

### 6.5.3 Mixed Methods

The combination of at least one qualitative and at least one quantitative component related to measurement scale, tools of data collection or data analysis technique in a single research study/project is known as mixed methods research. For example it can be in the following forms:

- Employment of more than one measurement procedure in different kind of situations such as different ways of measuring levels of job satisfaction.
- Employment of two or more methods of data collection. For example we may employ observation, interview and questionnaire to collect the data about decent work.
- Employment of two or more data analysis techniques for example, content analysis and factor analysis.

Mixed methods research has gained a tremendous popularity in social, behavioural and related sciences in recent years. The rationale for mixed method design research is to take the best of qualitative (QL) and quantitative methods (QN) and combine them. However, many debates on mixed method research design are based on methodological arguments due to the ways QN method is linked to positivist paradigm and QL method to Interpretative and critical theory paradigm.

As discussed above, keeping in view the diametrical oppositional attributes of QL and QN methods, one may argue that these methods are not compatible to each other and in such a situation how a truce between two kinds of approaches be useful to conduct the meaningful research?

Notwithstanding the philosophical debate between QN and QL approach. Mixed methods research designs are justified primarily by exploiting the

strength of each method and by combining their respective strengths within one single research design. The argument of incompatibility between two methods on philosophical ground is not valid because of the following reasons: (Bergman, 2008)

i) The QL and QN methods represent large and heterogeneous families of methods which vary within their own family to such an extent that it is difficult to identify unique set of qualities or attributes that encompasses the characteristics of one family of methods, clearly distinct from the characteristics of the members of the other family. Most characteristics encompass either only a subgroup of members of the family or are also applicable to some members of other family. Hence there is a need to re-think the division of labour between QL and QN methods in order to better understand the possibilities and functions of methods to better justify and apply mixed method design.

ii) The proposition of existence of one single reality under QN method is inconsistent with research applications. The emergent structures are based on co-construction between the researchers' selection and understanding of items in a questionnaire, their choice of analytic strategy and their interpretation of the statistical output on the one hand, and the way the respondents interpreted the survey questions with the given social, political, historical and economic context, on the other hand.

   From a methodological perspective, it is not proper to declare one approach more or less valid, valuable or scientific. Instead how to understand and analyze data need to be based to a large extent on the consistency formed between how to understand data in conjunction with the specific research question, rationale, aim etc. only in connection with the specificities of the research goals it makes sense to delimit the nature of reality. Thus, the decision on whether the researcher deals with one single reality, a constructed reality, multiple realities, multiple constructed realities or a co-constructed reality is unrelated to whether patterns in the data are detected via statistical analysis or otherwise.

iii) *A small vs large samples*: The sample size for many QN studies is often quite small and sample size for some QL studies can indeed be large.

iv) Regarding hypothesis testing, many researchers engaged in QL research, pursue conjectures that are embedded either in their research questions, the kind and the way they collect data, the way they analyse the data and the way, they protect themselves from selective reportings of their findings. In many types of well-established statical analyses, the QN methods like cluster analysis, factor analysis, correspondence analysis, multi-dimensional scaling are used in exploring data structures.

In order to avoid misconceptions and mistakes while deciding the research design issues, the following points need to be kept in mind.

a) We refrain ourselves from making specific argument for or against the QL and QN method for a specific research project. We need to focus our efforts more explicitly on embedding and justifying our selected methods according to our research questions, data needs, theoretical grounding and research design.

b) Data collection methods (i.e. unstructured narrative interview, survey research based on closed-ended questions) and data analysis methods (qualitative content analysis, discourse analysis, quantitative content analysis) should be differentiated.

c) We ought to be more aware of the actual inductive and deductive analytic phases of our research projects.

**Check Your Progress 2**

1) Which constituent of research study does guide the research design?

……

……

……

2) State the functions of research design?

……

……

……

3) What is the distinction between Mono Method and Mixed Method?

……

……

……

……

## 6.6 THE RATIONALE FOR MIXED METHODS RESEARCH

Mixed methods research has gained tremendous popularity in the social, behavioural and related sciences in recent years. This increasing popularity of mixed methods research is reflected in terms of the claims by mixed method researchers, increase in the publications on this topic and the inclusion of mixed methods designs in the text books. The argument for combining qualitative and quantitative methods is put forwarded on the following grounds.

i) The research methods associated with both quantitative and qualitative research have their own strengths and weaknesses. Combining them allows the researchers to offset their weakness to draw on the strengths of both.

ii) Mixed methods designs enable the researcher to provide a comprehensive account of the area of the enquiry in which he or she is interested if both quantitative and qualitative methods are used. for example, causal explanation of any event or phenomena can be captured by quantitative methods whereas the reason explanation and pattern explanation or the explanation involving processes can better be captured by qualitative methods.

iii) Quantitative and qualitative methods be used to answer different questions in a study. Further mixed methods can be utilized in order to gain complementary views about the same phenomenon or relationship. Research question for the two strands of the mixed study address related aspects of the same phenomenon.

iv) Mixed methods design can be used to develop the research further. Questions for one strand emerge from the inferences of a previous one (sequential mixed methods), or one strand provides hypothesis to be tested in the next one.

v) Mixed methods may also be used to assess the credibility of inferences obtained from one approach (strand). Exploratory and explanatory/ confirmatory questions in such situation, may be useful.

vi) Mixed methods are also used to obtain divergent pictures of the same phenomenon. These divergent findings can ideally be compared and contrasted.

vii) Quantitative research provides an account of structures in social life but qualitative research provides sense of process.

viii) Mixed methods research entails to generate hypothesis by using qualitative method and testing the hypothesis resorting quantitative method within a single project.

ix) Mixed methods research enables to present diversity of views i.e. combining researchers' and participants' perspectives through quantitative and qualitative research respectively by way of uncovering relationships between variables through quantitative research and revealing meaning among research participants through qualitative research.

## 6.7 FORM OF MIXED METHODS RESEARCH DESIGNS

Connecting qualitative and quantitative methods through sequencing and prioritising, there can be four forms of mixed methods research designs.

1) **Preliminary qualitative inputs to core quantitative method (qual-QUANT)**: With a view to use the strengths of qualitative methods to contribute to core quantitative methods, preliminary qualitative methods are used for collecting and interpreting the quantitative data. Preliminary qualitative methods can also be used as an input to generate hypotheses, develop content for questionnaires etc.

2) **Supplementary qualitative method as a follow up to core quantitative method (QUANT-qual)**: Supplementary qualitative method as a follow up to core quantitative method can be used to extend what is learned from quantitative study. It can demonstrate bases for results, examine reasons for failed hypotheses, explore the theoretical significance of outliers and so on.

3) **Preliminary quantitative inputs to core qualitative methods design (quan-QUAL)**: The study undertaken by using core quantitative method design gets inputs by way of using preliminary quantitative method for collecting and interpreting the qualitative data. Preliminary quantitative

method can also locate major differences between sub-groups, guide purposive sampling, establish results to pursue and so on.

4) **Supplementary quantitative method to core qualitative method design (QUAL-quan)**: Supplementary qualitative method as a follow up to core qualitative method can be used to develop measures of key concepts and relationships across different samples.

## 6.8 CASE STUDIES OF MIXED METHODS RESEARCH DESIGN

i) Studies where the quantitative component is dominant in terms of the framework of positivism and post positivism paradigm and inductive and deductive logic of enquiry guides the study, quantitative approach is more dominant and the researchers specialize in quantitative work. In such studies, qualitative data and their manipulation are shaped by the nature of research questions raised in the quantitative part of the project. The qualitative data are treated and transformed in effect into quantitative variables with the purpose of fleshing out the explanations required for the quantitative results. In such cases quantitative component follows qualitative component mixed methods. An illustration of a Research Study which has used Mixed Method pre dominated by the Quantitative Method is given below in the box:

**CASE 1: Mixed Method Research Design: Conducted the study within the framework of post positivist paradigm using inductive logic of inquiry and analyzing the qualitative data by quantitative technique of data analysis**

**Study: 'Decent Work: Concept, Measurement and Status in India': Ph.D thesis by Ms. Nausheen Nizami Ph.D Research Scholar, (Economics) IGNOU, N. Delhi**

The study was conducted within the framework of positivist paradigm using inductive logic of enquiry with the objectives to (i) examine the applicability of indicators of 'decent work in India (ii) develop indicators and tools to measure 'decent work (iii) evaluate the status of decent work in IT companies and to identify the socio-economic attributes influencing status of decent work, and (iv) know the employer's perspective in provision of decent work'.

**Methodology**: Decent work sums up the aspirations of the people in their professional life and is a revolutionary agenda of International Labour Office. Decent work is any such work which ensures provision of fair and free employment to all men and women of economically productive age-group in conditions of fairness, equity, security and dignity. It is a multidimensional concept that applies to both formal and informal sectors of an economy. The study is based on primary data collected through well designed questionnaire and telephonic interviews. Measurement of decent work was undertaken using both quantitative and qualitative methods of data analysis. A characteristic feature of this study was to transform the qualitative data in a manner so as to be able to use quantitative techniques of data analysis.

**Quantitative Techniques to analyse qualitative data**

a) Composite Index construction for Decent Work

With a view to measure the extent and range of decent work provision, decent work indices were constructed. These indices have been made for all the indicators of decent work used in this study by coding, normalising and aggregating the responses as the data was qualitative in nature. Thereafter, the standard formula developed by United Nation's Development Pr*o*gramme (UNDP) for the construction of HDI and other indices was applied.

b) Factor analysis of decent work indicators

The technique of factor-analysis was adopted to short-list the most relevant and reliable indicators of decent work. The principal component method was used to outline components (i.e. indicators in this study) explaining the maximum variance in the dependent variable.

c) Phi ϕ Coefficient of Correlation Analysis

Since the variables were dichotomous in nature, correlation analysis was attempted to test whether there was statistically significant associations between the responses on different decent work indicators. Partial correlation analysis examined the relationship between few socio-demographic variables and composite Decent work Index keeping the effect of other variables constant.

d) Chi-square test for independence

Chi-square test was conducted to test the independence of the results between the two scales of Decent work index (For example, 0 and 0-1) and results of each decent work indicator for those two scales.

**Thus an efficient mix of quantitative and qualitative methods were used to deal with the key issues related to decent work in this doctoral thesis.**

**Major Results:**

1) Composite decent work index reflects deficit in the status of decent work for a vast majority (94%) of IT employees.
2) Adequacy of earnings and productive employment are directly associated with decent work which is in consonance with the features of search and matching theory of employment.
3) Longer working hours have a significant impact on the physical and emotional well-being of an employee. A vast majority of IT employees (95%) were found to be working for longer hours.
4) Work-life balance and adequate earning and productive employment were found to be correlated. Majority of the employees in IT sector were found to be earning adequate earnings but their work-life balance was disrupted owing to several reasons including longer working hours. This implies that higher earnings of the employees in the IT industry are necessary

but not a sufficient condition for the employees. Their higher earnings are not resulting into their well-being (mental) as reflected in their deteriorating health status and imbalance between professional and personal lives.

5) Social security mechanisms, social dialogue and fair treatment at work places are major aspects of decent work owing to ensure sustainability of economic well-being of the employees.

6) Decent work reflects decent work place but a decent work place may not always translate into decent work because of the gap of the nature of work and work amenities and environment. This implies that deficit in decent work to the employees is correlated with deficit in decent work places.

7) Age and social class are the prominent attributes in influencing decent work status.

8) Despite the presence of premier educational institutions like IITs, RECs, etc. demand-supply gap in the available manpower for IT industry is the major concern of the employers pushing them to face the problem of adverse selection. A direct implication of this situation is the low employment elasticity of IT industry with respect to its growth and increasing contribution to GDP.

ii) Studies where qualitative component had priority in terms of the frame work of interpretative paradigm and abductive logic of enquiry. In such situations, the researchers identify themselves primarily as qualitative researchers. In such studies, the tools used for data/information collection is mixed in nature. The modalities of the different types of data are maintained and are not treated as commensurate. Integration of data takes place as a part of interpretative process referred to as analytic or interpretative integration. In such studies quantitative techniques can be used to analyse the perceptions of the people which may be qualitative in nature. An illustration of the research study using mix of qualitative and quantitative methods in the manner discussed here in is given below:

**CASE 2: Mixed Methods Research Design: Conducted the study within the framework of Qualitative Research Approach using Quantitative Methods for data analysis**

**Study: Gender specific impacts of climate variation in Agriculture:**

**By Mamta Mehar, Ph.D Research Scholar, IGNOU, N. Delhi**

The Main Objective of the study was to understand the gender differences in agriculture and analyse the linkages between gender and climate variation using mixed method approach. The research questions addressed in this study were: (i) Does there exist difference in gender perception on climate variability? (ii) Whether male and female farmers follow similar coping strategies to mitigate risk and shocks in agriculture? The study focused on to understand how gender and climate variation in agriculture are interrelated with the support of empirical evidences.

## Methodology

This study was conducted within the framework of qualitative approach anchored in interpretative paradigm using quantitative methods – questionnaire as a tool and survey as a method to collect the data and multivariate probit model to find out probability in deciding adaption strategy by male and female to climate variability.

This research study used case study method within a socio-cultural framework. Case study methodology derives much of its rationale and methods from ethnography characterized by "an instance in action" in a "real life context when the boundaries between the phenomenon and context are not clear". Keeping in view, the logic of enquiry, the research questions were analysed using both qualitative and quantitative methods. The qualitative analysis was done within the framework of "Interpretative paradigm of qualitative research". Since the research objective was to understand human behaviour within the surrounding context of climate change, where it was difficult to isolate the phenomena of climate variation and perception of males and females towards it, conducting of the present study under interpretive paradigm was appropriate. The associated research strategy used in this study was abduction. Abduction is the logic used to construct descriptions and explanations that are grounded in the everyday activities as well as in the language and meanings used by social actors. The social actors in the present proposal are men and women farmers, and their activities to be analysed are specific to decision making in climate change situation. The interpretative paradigm using abduction strategy was done via literature review as well as synthesis from findings of focus group discussion during the survey. In the wake of theoretical understanding, it was tried to quantify the impact using appropriate empirical analysis and using data from a baseline survey conducted by CCAFS in 2013 of 641 farm households in twelve villages of Vaishali districts in Bihar, India; it was attempted to quantify the gender differentiated impact.

## Results

The linkage between gender and climate variation in agriculture was more influenced from subjective experiences of individuals or society. The movement of understanding "is constantly from the whole to the part and back to the whole" (Gadamer, 1976). According to Gadamer, it is a circular relationship. Since the study aims to understand the perception of human in context of climate variation, it cannot be explained entirely by any scientific paradigm where hypothesis are pre-structured and to some extent results are known. The interpretative paradigm points out that positivism, in its first attempt to model the social after the natural sciences, fails to see that unlike nature, social reality exists only insofar as lay members create that reality in meaningful interaction (Fuchs S., 1992: 198). But the point of interpretive paradigm is that we must first understand social worlds from within, before we develop scientific models and explanations (Fuchs, S. 1992: 205). According to Willis (1995) interpretivists are anti-foundationalists, who believe there is no single correct route or particular method to knowledge. Walsham (1995) argues that in the interpretive tradition there are no 'correct' or 'incorrect' theories. Instead, they should be judged according to how 'interesting' they are to the researcher as well as those involved in the same areas.

The linkages between gender and climate variation in context of agriculture are bi-directional. Changes in climate event effect agriculture through reduction in yield, increase in occurrence of pests or crop disease, changes in time of activities for example, changes in winter time could result in change in harvesting or sowing days of respective crops. All this affects farmers' livelihood and work activity in different ways. For example, women (and children) being the ones traditionally charged with water collection roles, have to travel longer distances after drought. Evidences suggest decrease in yield and thereby total availability of food consumption in a household basket usually has been seen as complemented by reduction in meals of female. This exposes them to health issues. Male farmers are usually seen to migrate in search of nonfarm opportunities outside their village domain, putting all burden or responsibilities on women. Moreover mostly in rural areas women are less likely to get education as compare to male, they are less allowed to mobile and do more household related activities, and they have less control over the family assets.

The other direction of gender differentiated impacts in context of climate events in agriculture can be viewed through understanding their coping strategies. The adaption of particular strategies are often influenced by the priorities and options available for women and men for coping these strategies. Constraints on women's mobility, and behavioural restrictions hinders their ability to make decision on various matters and thus they have less choices of coping strategies available. Additionally inequitable access to assets and resources such as land, knowledge, technology, limited power in decision-making, education, health care and food have been identified as determinant factors behind adaptation of different coping strategies of male versus female. The results of quantitative analysis also supports this argument.

The females are mostly illiterate (71 per cent) and have little access to resources. Only 9 per cent of farm household have reported to provide land entitlement to at least one female member. Additionally the gender power dynamics is seen to favour only one gender i.e. male. The decision making role in different household related expenditures as well as in agriculture activities is dominated by male. Considering the perception of male vis-à-vis female farmer about effects of climate change has shown a similar pattern. More than 90 per cent of farmers are aware about climate variation and feels that weather conditions adversely affect their agriculture activities. However again, the decisions to cope with the risk due to climate have suggested different behaviour, priorities by male and female. Survey results suggest that almost 60 per cent of surveyed farmers have adapted at least one coping strategy and in total adoption of more than 30 coping strategies are reported. The results of multivariate probit model suggest that male farmers have higher probability in deciding adaptations through additional jobs as well as specific agriculture strategies such as crop rotation and use of hybrid seed in farming practices.

iii) Studies presenting diversity of views i.e. combining researcher's and participants' perspectives through quantitative and qualitative research. Researchers uncover relationship between variables through quantitative methods and reveal meaning among research participants through qualitative methods. Thus a comprehensive account of the area of the enquiry is presented. An illustration of the study using this form of mixed method research is given below:

**CASE 3: Mixed Method Research –used Quantitative Methods for data analysis and Qualitative Methods for revealing and understanding participants' perspective on quality of education**

**Study: 'Measuring Quality of Education and Its Determinants: Indian Context': Ph.D thesis submitted by Ms. Charu Jain, Ph.D Research Scholar, (Economics) IGNOU, New Delhi**

The study was carried out with the objectives to (i) examine the linkages of secondary education with various socio-economic outcomes at all India level and state level, (ii) identify various determinants (including students' family background characteristics, and schools' characteristics) affecting students' performances and teaching efficiencies in senior secondary schools in Delhi. To fulfill these objectives, both quantitative and qualitative research approaches were applied.

**Quantitative Method:** The quantitative approach was used to test the hypothesis generated. Two set of questionnaires one related to the teachers and the other to students were got filled up through in-depth face to face interviews with the respondents. The student questionnaire was translated in Hindi language to enhance the understanding level of students in Hindi medium government schools in Delhi. Within translation procedure, the student questionnaire was first translated in Hindi and thereafter re-translated in English to re-check the accuracy and meaning of questions. The teachers' questionnaire was provided in English language only. The information thus retrieved from these questionnaires was used to measure the performance of students and teacher analyzing thereby the overall quality of education at lower secondary and senior secondary levels of education. The information collected through quantitative approach was statistically analyzed and tested using various statistical softwares like SPSS and STATA?

**Qualitative Method:** In the qualitative approach, direct observation method and key informer interviews were used to supplement the information/data collected through quantitative method by way of two sets of structured questionnaires. Qualitative research method enabled the researcher to approach the research topic from the perspective of teachers and students. Further, the qualitative techniques were used to examine social processes which otherwise could not have been captured by traditional quantitative measures.

*The direct observation technique* enabled researcher to learn about the behaviour of the people under study: students and teachers in the natural setting i.e. schools through observing. It gave the researcher the first hand information on the quality aspects of school and classrooms and working conditions for teachers, which were quite useful in doing the analysis.

The field notes, which were taken down during the period of fieldwork, were instrumental during analysis and interpretation of the results, as they helped in recalling the incidents that took place at the time of data collection. *Face-to-face interviews* with the key informants were also conducted within qualitative approach. Key informant interviews were qualitative informal interviews with people who knew what was going on in the community. Our purpose of conducting key informant interviews was to collect information from people who had firsthand knowledge about the overall functioning of school. The key informers in the sample schools were selected using purposive sampling technique and in this case the key informants were either senior teacher in school or senior administrative person in schools. Interviewing key informants enabled the researcher to look at the underlying issues and problems with varying perspectives. These interviews were conducted using an informal approach. Initially the questions were drafted in a form of well designed open ended short questionnaire for conducting face to face interviews with the respondents. One respondent or a group of 2-3 people from each school were selected purposively for gathering this information, depending upon how well they were informed about the school.

Apart from these informal interviews, the school guards/receptionists were also contacted informally to gather the information on schools. The information thus collected through qualitative approach was first entered in Excel in a pre-designed format and then coded so that it can be further analysed statistically. Moreover, the responses to few open ended questions were reviewed and inferences were drawn which were quite useful in providing suggestions and recommendation for this study.

**Main Results**

1) Secondary education increases income level of individuals to a significant extent. One unit increase in Gross Enrollment Ratio (GER) at secondary level leads to increase in the personal disposable income by 0.65 units.
2) The educational background of households determine their occupational pattern. The households with at least secondary education are either salaried or business class, while those with elementary level of education are more concentrated in agricultural/wage earning activities. Higher the level of education among females, greater the proportion of female work participation rate.
3) Secondary education bears highest impact on improving health indicators. Secondary education has high potential for bringing demographic transitions in terms of decline in maternal mortality rates, infant mortality rates, fertility rates, death rates etc.
4) Secondary education attainment brings behavioral changes in individuals in terms of higher savings and diversion in the expenditure pattern.
5) School resources, teachers and teaching quality, student's family background, mass media exposure and self motivational factors of students determine their outcome to significant extent.
6) Within school characteristics, cleanliness in school, well qualified teachers their friendly and supportive attitude with students and quality of school infrastructure influence students performance positively. Lack of student motivation, large class size, lack of school resources and facilities, lack of adequate teaching material affect the teaching abilities adversely.

**Check Your Progress 3**

1) Give two situations where Mixed Method research is superior to Mono Method?

……

……

……

……

2) Give an example of Mixed Method where qualitative method pre-dominates quantitative method?

……

……

……

……

3) State the different forms of Mixed Methods?

……

……

……

……

## 6.9 LET US SUM UP

Traditionally two distinct research approaches – Quantitative (QN) and Qualitative (QL) – have been followed in undertaking research in social sciences. The former is associated with positivism and post positivism paradigm and the latter principally with interpretative paradigm and critical theory paradigm. Depending upon the nature and field of specialization research studies may be of various types. Important among these are – theoretical and applied, descriptive and explanatory, conceptual and empirical, exploratory and experimental etc.

Research design is a logical structure of enquiry specifying the types of evidences needed to answer the research questions. Research methods are made of tools for data collection, type of data, sampling design, sample size, techniques for data analysis etc. There is nothing intransic about any research design that requires a particular method of data collection. Research Methods can broadly be of two types – Mono Methods and Mixed Methods. Within mono methods, it can be either quantitative or qualitative. The combination of at least one qualitative and at least one quantitative component in a single research project is known as mixed methods research. The mixed methods have gained popularity in carrying out research in social and behavioral sciences because combining them allows the researchers (i) to offset their (QN and QL methods) weaknesses, (ii) provide comprehensive account of the area of the enquiry, (iii) to answer different types of questions. Studies undertaken under mixed methods can be of four types – (i) preliminary quantitative methods

input design and core qualitative methods as core design (quan-QUAL), (ii) supplementary quantitative methods design as follow up to core quantitative methods design (QUAL-quan), (iii) preliminary qualitative methods input design and core quantitative methods design (qual-QUAN), (iv) supplementary qualitative methods design as follow up to quantitative methods design (QUAN-qual).

## 6.10 SOME USEFUL BOOKS

Bergman Manfred Max (2008): *The Strawmen of Qualitative-Quantitative Divide and Their Influence on Mixed Method Research in Advances in Mixed Methods Research*, edited by Manfred Max Bergman, Sage Publications Ltd. London.

Creswell John W, Clark Kicki L Plano and Amanda L. Garrett (2008): *Methodological Issues in Conducting Mixed Methods Research Designs in Advances in Mixed Methods Research*, Sage Publications Ltd. London.

Tashakkori Abbas and Teddlie Charles (2008): *Quality of inferences in Mixed Methods Research in Advances in Mixed Methods Research*, Sage Publications Ltd. London.

Morgan David L (2014): *Integrating Qualitative and Quantitative Methods- A Pragmatic Approach,* Sage Publications, London, N. Delhi, Singapore.

## 6.11 ANSWERS OF HINTS TO CHECK YOUR PROGRESS EXERCISES

**Check Your Progress 1**

1) A variable expressed in terms of numerals like 1, 2, 3,.. etc. is called quantitative variable. For instance, yield of an agricultural output, height/weight of individuals, etc. can be measured and expressed in quantitative terms. On the other hand, a variable expressed in terms of the choices or opinion about attributes/ characteristics like one's feelings or opinion is called qualitative variable.

2) See Sub-section 6.2.1

3) See Sub-section 6.2.3

4) Descriptive research describes a situation, a phenomenon or event or a social system. It aims to describe the state of affairs as it exists. On the other hand, explanatory results aim to establish cause and effect relationship.

**Check Your Progress 2**

1) Research questions guide the research design.

2) There are two major functions of the research design: (i) to determine the type of evidences (data/information) which are needed to answer the research questions, (ii) to eliminate the alternative explanation of the results.

3) If a research study is confined to use **either** quantitative **or** qualitative method, it is said to be a mono method research. The combination of at least one quantitative and at least one qualitative component in a single research project is known as mixed methods research.

**Check Your Progress 3**

1) See Section 6.6
2) See Section 6.7
3) See Section 6.7

# UNIT 7 DATA COLLECTION AND SAMPLING DESIGN

**Structure**



## 7.0 OBJECTIVES

After going through this Unit, you will be able to:

- appreciate different methods of collecting data;
- acquire knowledge of different tools of data collection;
- define the key terms commonly used in quantitative analysis, like parameter, statistic, estimator, estimate, inference, standard error, confidence intervals, etc.;
- distinguish between random and non-random sampling procedures for data collection;
- appreciate the advantages of random sampling in the assessment of the "precision" of the estimates of population parameters;
- acquire knowledge of the procedure for drawing samples by different methods;

- develop the ability to obtain estimates of key parameters like population and sample aggregates, proportion, mean, etc.; and of the "precision" of such estimates under different sampling methods; and
- appreciate the feasibility/appropriateness of applying different sampling methods in different research contexts.

## 7.1 INTRODUCTION

Studies of the behaviour of variables and of relationships amongst them necessitate measurement of the variables involved. Variables can be *quantitative variables* like GNP or *qualitative variables* like opinions of individuals on, say, ban on smoking in public places. The former set assumes quantitative values. The latter set does not admit of easy quantification, though some of these can be categorised into groups that can then be assigned quantitative values. Research strategies thus adopt two approaches, quantitative and qualitative. We shall deal with the quantitative approach in this unit. The various measurement scales and scaling techniques to measure the qualitative data will be discussed in Unit 8.

The basic ingredient of quantitative research is the measurement in quantitative terms of the variables involved or the collection of the data relevant for the analytical and interpretative processes that constitute research. The quality of the data utilised in research is important because the use of faulty data in such endeavour results in misleading conclusions, however sophisticated may be the analytical tools used for analysis. Research processes, be it testing of hypotheses and models or providing the theoretical basis for policy or review of policy, call for objectivity, integrity and analytical rigour in order to ensure academic and professional acceptability and, above all, an effective tool to tackle the problem at hand. Data used for research should, therefore, reflect, as accurately as possible, the phenomena these seek to measure and be free from errors, bias and subjectivity. Collection of data has thus to be made on a scientific basis.

## 7.2 AN OVERVIEW OF THE UNIT

How to assemble data on a scientific basis? There are broadly three different methods of collecting data. These are dealt with in section 7.3. The tools that one can use for collecting data – the formats and the devices that modern technology has provided – are enumerated in section 7.4. There are situations where it is considered desirable to gather data from only *a part of* the universe, or a *sample* selected from the universe of interest to the study at hand, rather than a complete coverage of the universe, for reasons of cost, convenience, expediency, speed and effort. Questions then arise as to the manner in which such a *sample* should be chosen – the sampling design. This question is examined in detail in section 7.5. The discussion is divided into a number of sub-topics. Concepts relating to population aggregates like mean and variance and similar aggregates from the sample and the use of the latter as estimates of population aggregates have been introduced in sub-section 7.5.1. There are two types of sampling: random and non random. Non random sampling methods and the contexts in which these are used are described in sub-section 7.5.2. A random sample has certain advantages over a non random sample — it provides a basis for drawing *valid* conclusions from the sample about the parent population. It enables us to state the precision of the estimates of

population parameters in terms of (a) the extent of their variation or (b) an interval within which the value of the population parameter is likely to lie with a given degree of certainty. Further, it even helps the researcher to determine the size of the sample to be drawn if his project is subject to a sanctioned budget and permissible limits of error in the estimate of the population parameter. These principles are explained in sub-section 7.5.3. Eight methods of random sampling are then detailed in sub-section 7.5.4. These details relate to (i) operational procedures for drawing samples, and (ii) expressions for (a) estimators of parameters and measures of their variation and b) estimators of such variation where the population variation parameter is not known. Different sampling procedures are also compared, as we go along, in terms of the relative precision of the estimates, they generate. Finally, the question of choosing the sampling method that is appropriate to a given research context is addressed in section 7.6. A short summing up of the Unit is given in section 7.7.

## 7.3 METHOD OF DATA COLLECTION

There are three methods of data collection – *the Census and Survey Method, the Observation Method and the Experimental Method.* The first is a carefully planned and organised study or enquiry to collect data on the subject of the study/enquiry. We might for instance organise a study on the prevalence of the smoking habit among high school children – those aged 14 to 17 - in a certain city. One approach is to collect data of the kind we wish to collect on the subject matter of the study from *all* such children in *all* the schools in the city. In other words, we have a ***complete enumeration*** or ***census*** of the ***population*** or ***universe*** relevant to the enquiry, namely, the city's high school children (called the *respondent units or informants* of the Study) to collect the data we desire. The other is to confine our attention to a *suitably selected part of the population* of high school children of the city, or a ***sample,*** for gathering the data needed. We are then conducting a *sample survey.* A well known example of Census enquiry is the **Census of Population** conducted in the year **2011,** where data on the demographic, economic, social and cultural characteristics of *all persons* residing in India were collected. Among **sample surveys** of note are the household surveys conducted by the National Sample Survey Organisation (NSSO) of the Government of India that collect data on the socio-economic characteristics of a sample of households spread across the country.

*The Observation Method* records data as things occur, making use of an appropriate and accepted method of measurement. An example is to record the body temperature of a patient every hour or a patient's blood pressure, pulse rate, blood sugar levels or the lipid profile at specified intervals. Other examples are the daily recording of a location's maximum and minimum temperatures, rainfall during the South West / North East monsoon every year in an area, etc.

*The Experimental Method* collects data through *well designed and controlled* statistical experiments. Suppose for example, we wish to know the rate at which manure is to be applied to crops to maximise yield. This calls for an experiment, in which all variables other than manure that affect yield, like water, quality of soil, quality of seed, use of insecticides and so on, need to be controlled so as to evaluate the effect of different levels of manure on the yield. Other methods of conducting the experiment to achieve the same objective without controlling "all other factors" also exist. Two branches of statistics – The Design and Analysis of Experiments and Analysis of Variance — deal with these.

## 7.4 TOOLS OF DATA COLLECTION

How do we collect data? We translate the data requirements of the proposed Study into items of information to be collected from the respondent units to be covered by the study and organise the items into a logical format. Such a format, setting out the items of information to be collected from the respondent units, is called ***the questionnaire or schedule*** of the study. The questionnaire has a set of pre-specified questions and the replies to these are recorded either by the respondents themselves or by the investigators. The *questionnaire approach* assumes that the respondent is capable of understanding and answering the questions all by himself/herself, as the investigator is not supposed, in this approach, to influence the response in any manner by interpreting the terms used in the questions. Respondent-bias will have to be minimised by keeping the questions simple and direct. Often the responses are sought in the form of "yes", "no" or "can't say" or the judgment of the respondent with reference to the perceived quality of a service is graded, like, "good", "satisfactory" or "unsatisfactory".

In the *schedule approach* on the other hand, the questions are detailed. *The exact form of the question to be asked of the respondent is not given to the respondent and the task of asking and eliciting the information required in the schedule is left to the investigator.* Backed by his training and the instructions given to him, the investigator uses his ingenuity in explaining the concepts and definitions to respondents to obtain reliable information. This does not mean that investigator-bias is more in the schedule approach than in the questionnaire approach. Intensive training of investigators is necessary to ensure that such a bias does not affect the responses from respondents.

Schedules and questionnaires are used for collecting data in a number of ways. Data may be collected by personally contacting the respondents of the survey. Interviews can also be conducted over the telephone and the responses of the respondent recorded by the investigator. The advent of modern electronic and telecommunications technology enables interviews being done through e-mails or by 'chatting' over the internet. The mail method is one where (usually) questionnaires are mailed to the respondents of the survey and replies received by mail through (postage pre-paid) business-reply envelopes. The respondents can also be asked (usually by radio or television channels or even print media) to send their replies by SMS to a mobile telephone number or to an e-mail address.

Collection of data can also be done through mechanical, electro-mechanical or electronic devices. Data on arrival and departure times of workers are obtained through a mechanical device. The time taken by a product to roll off the assembly line and the time taken by it to pass through different work stations are recorded by timers. A large number of instruments are used for collecting data on weather conditions by meteorological centres across the country that help assessing current and emerging weather conditions. Electronic Data Transfers (EDT) can also be the means through which source agencies like ports and customs houses, where export and import data originate, supply data to a central agency like the Directorate General of Commercial Intelligence and Statistics (DGCI&S) for consolidation.

The above methods enable us to collect *primary data,* that is, data being *collected afresh* by the agency conducting the enquiry or study. . The agency

concerned can also make use of data on the subject *already collected* by another agency or other agencies – *secondary data.* Secondary data are published by several agencies, mostly Government agencies, at regular intervals. These can be collected from the publications / compact discs or the websites of the agencies concerned. But such data have to be examined carefully to see whether these are suitable or not for the study at hand before deciding to collect new data.

*Errors* in data constitute an important area of concern to data users. Errors can arise due to confining data collection to a sample. (*sampling errors*). It can be due to faulty measurement arising out of lack of clarity about what is to be measured and how it is measured. Even when these are clear, errors can creep in due to inaccurate measurement. Investigator bias also leads to errors in data. Failure to collect data from respondent units of the population or the sample due to omission by the investigator or due to non-response (respondents not furnishing the required information) also results in errors. (*non-sampling errors*). The *total survey error* made up of these two types of errors need to be minimised to ensure quality of data.

## 7.5 SAMPLING DESIGN

We have looked at methods and tools of data collection, chief among which is the sample survey. How to select a sample for the survey to be conducted? There are a number of methods of choosing a sample from a universe. These consist of two categories, *random sampling* and *non-random sampling*. Let us turn these methods and see how well the results from the sample can be utilised to draw conclusions about the parent universe.

But first let us turn to some notations, concepts and definitions.

### 7.5.1 Population and Sample Aggregates and Inference

Let us denote population characteristics by upper case (capital) letters in English or Greek and sample characteristics by lower case (small) letters in English. Let us consider a (finite) population consisting of N units $U_i$ (i = 1,2,….N). Let $Y_i$ (i = 1,2,….N) be the value of the variable y, the characteristic under study, for the $i^{th}$ unit $U_i$ (i = 1,2,…..N). For instance, the units may be the students of a university and y may be their weight in kilograms. Any function of the population values $Y_i$ is called a ***parameter***. An example is the population mean ‘μ’ or ‘M’ given by $(1/N) \sum_i Y_i$, where $\sum_i$ stands for summation over i = 1 to N. Let us now draw a sample of ‘n’ units $u_i$ (i = 1,2,…..n)[1] from the above population and let the value of the $i^{th}$ sample unit be $y_i$ (i = 1,2,…..n)[2]. In other words, $y_i$ (i = 1,2,….n) are the sample observations. A function of the sample observations is referred to as a ***statistic***. The sample mean ‘m’ given by $(1/n) \sum_i y_i$, $\sum_i$ (i =1 to n), is an example of a statistic.

Let us note the formulae for some important parameters and statistics.

Population total $Y = \sum_i Y_i$, $\sum_i$ stands for summation over i = 1 to N (2.1)

Population mean = ‘μ’ or ‘M’ , $= (1/N) \sum_i Y_i$, $\sum_i$ , i = 1 to N (2.2)

[1] The sample units are being referred to as $u_i$ (i = 1,2,…..n) and not in terms of $U_i$ as we do not know which of the population units have got included in the sample. Each $u_i$ in the sample is some population unit

[2] The same reasons apply for referring the sample values or observations as $y_i$ (i = 1,2,…n) and not in terms of the population values $Y_i$ . $y_i$ wil be some $Y_i$.

Population variance $\sigma^2 = (1/N) \sum_i Y_i^2 - M^2$, $\sum_i$, i = 1 to N (2.3)

Population SD = $\sigma = +\sqrt{[(1/N) \sum_i Y_i^2 - M^2]}$, $\sum_i$, i = 1 to N (2.4)

Sample mean = $(1/n) \sum_i y_i$, $\sum_i$, (i =1 to n) (2.5)

Sample variance $s^2 = (1/n) \sum_i y_i^2 - m^2$, $\sum_i$ i = 1 to n (2.6)

$= [ss]^2 /n$, where $[ss]^2 = \sum (y_i - m)^2 = \sum y_i^2 - n m^2 =$ sum of squares of sample observations from their mean 'm' (2.7)

Sample standard deviation 's' = $+\sqrt{[(1/n) \sum_i y_i^2 - m^2]}$, $\sum_i$, i = 1 to n (2.8)

Population proportion $P = (1/N) \sum_i Y_i = N^1 / N$, (where $N^1$ is the number of units in the population possessing a specified characteristic) (2.9)

$\sigma^2 = (1/N) \sum_i Y_i^2 - M^2 = P - P^2 = P(1 - P) = PQ$,

where $Q = [(N - N^1)/N] = (1 - P)$. (2.10)

m = p, (proportion of units in the sample with the specific characteristic) (2.11)

$s^2 = p(1 - p) = pq$, where p is the sample proportion and p + q = 1 (2.12)

$[ss]^2 = npq$ (2.13)

The purpose of drawing a sample from a population is to arrive at some conclusions about the parent population from the results of the sample. This process of drawing conclusions or making inferences about the population from the information contained in a sample chosen from the population is called ***inference.*** Let us see how this process works and what its components are. The sample mean 'm', for example, can serve as an estimate of the value of the population mean 'μ'. The statistic 'm' is called an ***estimator (point estimator)*** of the population mean 'μ'. The value of 'm' calculated from a specific sample is called an ***estimate (point estimate)*** of the population mean 'μ'. In general, *a function of sample observations, that is, a statistic, which can be used to estimate the unknown value of a population parameter, is an* ***estimator*** *of the population parameter. The value of the estimator calculated from a specific sample is an* ***estimate*** *of the population parameter.*

The estimate '$m_1$' of the population parameter 'μ', computed from a sample, will most likely be different from 'μ'. There is thus an *error* in using '$m_1$' as an estimate of 'μ'. This error is the sampling error, assuming that all measurement errors, biases etc., are absent, that is, *there are no non-sampling errors.* Let us draw another sample from the population and compute the estimate '$m_2$' of 'μ'. '$m_2$' may be different from '$m_1$' and also from 'μ'. Supposing we generate in this manner a number of estimates $m_i$ (i = 1,2,3,…….) of 'μ' by drawing repeated samples from the population. All these $m_i$ (i = 1,2,3,….) would be different from each other and from 'μ'. What is the extent of the variability in the $m_i$ (i = 1,2,3,….), or, the variability of the error in the estimate of 'μ' computed from different samples? How will these values be spread or scattered around the value of 'μ' or the errors be scattered around zero? What can we say about the estimate of the parameter obtained from the specific sample that we have drawn from the population as a means of measuring the parameter, without actually drawing repeated samples? How well do non-random and random samples answer these questions? The answers to these questions are important from the point of view of inference.

Let us first look at the different methods of non-random sampling and then move on to random sampling.

## 7.5.2 Non-Random Sampling

There are several kinds of non-random sampling. **A judgment sample** is a sample that has been selected by making use of one's expert knowledge of the population or the universe under consideration. It can be useful in some circumstances. An auditor for example could decide, on the basis of his experience, on what kind of transactions of an institution he would examine so as to draw conclusions about the quality of financial management of an institution. **Convenience Sampling** is used in exploratory research to get a broad idea of the characteristic under investigation. An example is one that consists of some of those coming out of a movie theatre; and these persons may be asked to give their opinion of the movie they had just seen. Another example is one consisting of those passers by in a shopping mall whom the investigator is able to meet. They may be asked to give their opinion on a certain television programme. The point here is the convenience of the researcher in choosing the sample. **Purposive Sampling** is much similar to judgement sampling and is also made use of in preliminary research. Such a sample is one that is made up of a group of people specially picked up for a given purpose. **In Quota Sampling**, subgroups or strata of the universe (and their shares in the universe) are identified. **A convenience** or **a judgement sample** is then selected from each stratum. No effort is made in these types of sampling to contact members of the universe who are difficult to reach. In Heterogeneity Sampling units are chosen to include all opinions or views. **Snowball Sampling** is used when dealing with a rare characteristic. In such cases, contacting respondent units would be difficult and costly. This method relies on referrals from initial respondents to generate additional respondents. This technique enables one to access social groups that are relatively invisible and vulnerable. This method can lower search costs substantially but this saving in cost is at the expense of the representative character of the sample. An example of this method of sampling is to find a rare genetic trait in a person and to start tracing his lineage to understand the origin, inheritance and etiology of the disease.

It would be evident from the description of the methods given above that the relationship between the sample and the parent universe not clear. The selection of specific units for inclusion in the sample seem to be *subjective* and *discretionary* in nature and, therefore, may well reflect the researcher's or the investigator's attitudes and bias with reference to the subject of the enquiry.

A sample has to be *representative* of the population from which it has been selected, if it is to be useful in arriving at conclusions about the parent population. ***A representative sample*** *is one that contains the relevant characteristics of the population in the same proportion as in the population.* Seen from this angle, the non-random sampling methods described above do not yield representative samples. Such samples are, therefore, not helpful in drawing *valid* conclusions about the parent population and *the way these conclusions change* when another sample is chosen from the population. Non-random sampling is, however, useful in certain circumstances. For instance, it is an inexpensive and quick way to get a preliminary idea of the variable under study or a rough preliminary estimate of the characteristics of the universe that helps us to design a scientific enquiry into the problem later. It is thus useful in exploratory research.

**Check Your Progress 1**

1) Name the three methods of the data collection

……

……

……

2) What is meant by a 'parameter'? Defining the term 'statistic', indicate the expressions for population/sample mean and population/sample mean and population/sample variance.

……

……

……

3) Mention the most important purpose of studying the population on the basis of a sample. In this context, define the terms 'estimator' and 'estimate' with a suitable example.

……

……

……

……

4) Defining the term 'representative sample'. Indicate how is 'random sampling' principally different from that of 'non-random sampling'? What could be the use of the latter despite its major drawback vis-à-vis the former?

……

……

……

……

## 7.5.3 Random or Probability Sampling

Random sampling methods, on the other hand, yield samples that are representative of the parent universe. The selection process in random sampling is free form the bias of the individuals involved in drawing the sample as the units of the population are selected at random for inclusion in the sample. *Random sampling is a method of sampling in which each unit in the population has a predetermined chance (probability) of being included in the sample. A sampling design is a clear specification of all possible samples of a given type with their corresponding probabilities.* This property of random sampling helps us to answer the questions we raised at the end of sub-section 2.6.1 above. That is, we can make estimates of the characteristics of the parent population from the results of a sample and also indicate the extent of error to which such estimates are subject or the *precision of the estimate*. This is better than not knowing anything at all about the magnitude of the error in our statements regarding the parent population. Let us see how random sampling helps in this regard.

## A) Precision of Estimates – Standard Errors and Confidence Intervals

We noted earlier (the last paragraph of sub-section 7.5.1) that the sample mean (an estimate of the population mean 'μ') will have different values in repeated samples drawn from the population and none of these may be equal to 'μ'. Suppose that the repeated samples drawn from the population are *random* samples. The sample mean computed from a *random sample* is a *random variable*. So is the sampling error, that is, the difference between'μ' and the sample mean. The values of the sample means (and the corresponding errors in the estimate of 'μ') computed from the repeated random samples drawn from the population are the values assumed by this random variable with probabilities associated with drawing the corresponding samples. These will trace out a frequency distribution that will approach a probability distribution when the *number* of random samples drawn increases indefinitely. *The probability distribution of sample means computed from all possible random samples from the population is called the* ***sampling distribution of the sample mean***. The sampling distribution of the sample mean has a mean and a standard deviation. *The sample mean is said to be an* ***unbiased estimator*** *of the population mean if the mean of the sampling distribution of the sample mean is equal to the mean of the parent population, say, μ. In general, an estimator "t" of a population parameter "θ" is an unbiased estimator of "θ" if the mean of the sampling distribution of "t", or the expected value of the random variable "t", is equal to "θ"*. In other words, the mean of the estimates of the parameter made from all possible samples drawn from the population will be equal to the value of the parameter. Otherwise, it is said to be a ***biased estimate***. Supposing the mean of the sampling distribution of sample mean is Kμ or K+μ, where K is a constant. The bias in the estimate can be easily corrected in such cases by adopting m/K or (m – K) as the estimator of the population mean.

The variance of the sampling distribution of the sample mean is called the ***sampling variance of the sample mean***. The standard deviation of the sampling distribution of sample means is called ***the standard error (SE) of the sample mean***. It is also called the ***standard error of the estimator (of the population mean)***, as the sample mean is an estimator of the population mean. *The standard error of the sample mean is a measure of the variability of the sample mean about the population mean or a measure of the* ***precision*** *of the sample mean as an estimator of the population mean.* The ratio of the standard deviation of the sampling distribution of sample means and the mean of the sampling distribution is called the ***coefficient of variation (CV)*** of the sample mean or the ***relative standard error (RSE)*** of the sample mean. That is,

$$\text{CV or RSE} = C = \text{standard deviation / mean} \qquad (2.14)$$

CV (or RSE) is a free number or is dimension-less, while the mean and the standard deviation are in the same units as the variable 'y'. (These definitions can easily be generalised to the sampling distribution of any sample statistic and its SE and RSE.)

We have talked about the unbiasedness and precision of the estimate made from the sample. What more can we say about the precision of the estimate and other characteristics of the estimate? *This is possible if we know the nature of the sampling distribution of the estimate.*

The nature of the sampling distribution of, say, the sample mean, or for that matter any statistic, depends on the nature of the population from which the random sample is drawn. If the parent population has a normal distribution with mean $\mu$ and variance $\sigma^2$ or, in short notation, N ($\mu$, $\sigma^2$), the sampling distribution of the sample mean, based on a random sample drawn from this, is N ($\mu$, $\sigma^2/n$). In other words, the variability of the sample mean is much smaller than that of the variable of the population and it also *decreases* as the sample size *increases*. ***Thus, the precision of the sample mean as an estimate of the population mean increases as the sample size increases***.

As we know, the normal distribution N ($\mu$, $\sigma^2$) has the following properties:

i) Approximately 68% of all the values in a normally distributed population lie within a distance of one standard deviation (plus and minus) from the mean,

ii) Approximately 95% of all the values in a normally distributed population lie within a distance of 1.96 standard deviation (plus and minus) of the mean,

iii) Approximately 99% of all the values in a normally distributed population lie within a distance of 2.576 standard deviation (plus and minus) of the mean.

The statement at (iii) above, for instance, is equivalent to saying that the population mean $\mu$ will lie between the observed values ($y - 2.576\ \sigma$) and ($y + 2.576\ \sigma$) in 99% of the random samples drawn from the population N($\mu$, $\sigma^2$). Applying this to the sampling distribution of the sample mean, which is N($\mu$, $\sigma^2/n$), we can say that

$$\Pr.[(m - 2.576\ \sigma / \sqrt{n}) < \mu < (m + 2.576\ \sigma / \sqrt{n})] = 0.99 \qquad (2.15)$$

or that the population mean $\mu$ will lie between the limits computed from the sample, namely, ($m - 2.576\ \sigma/\sqrt{n}$) and ($m + 2.576\ \sigma\sqrt{n}$) in 99% of the samples drawn from the population. This is an ***interval estimate,*** or a ***confidence interval,*** for the parameter with a ***confidence coefficient of 99%*** derived from the sample.

***The general rule for constructing a confidence interval of the population mean with a confidence coefficient of 99% is: the lower limit of the confidence interval is given by the "estimate of the population mean minus 2.576 times the standard error of the estimate" and the upper limit of the interval by the "estimate plus 2.576 times the standard error of the estimate".*** (2.16)

B) **Assessment of Precision – Unknown Population Variance**

If the parent population is distributed as N (M, $\sigma^2$) and $\sigma^2$ is not known, we make use of an estimate of $\sigma^2$. The statistic '$s^2$ ' given in formula *2.6* can be one such, but this is not an unbiased estimate of $\sigma^2$ as E ($s^2$) = [(n – 1)/ n] $\sigma^2$. We, therefore, by using (2.6) and (2.7) have:

$v(y)=[ns^2/(n-1)]$ as an unbiased estimate of $\sigma^2$. (2.17)

$v(y) = [1/(n-1)]\ [ss]^2$ or, $v(y) = [(1/(n-1))[\ \ _i y_i^{\ 2} - nm^2\ ]$ (2.18)

As the sampling variance of the sample mean 'm' is $\sigma^2/n$, *an unbiased estimate v(m) of the sampling variance will be v(y)/n.* Let us now consider the statistic defined by the ratio,

$$t = (m - M) / [\sqrt{v(y)} / \sqrt{n}] \quad (2.19)$$

The numerator is a random variable distributed as $N(0, \sigma^2/n)$ and the denominator is the square root of the *unbiased estimate of its variance*. The sampling distribution of the statistic 't' is the Student's t-distribution with (n – 1) degrees of freedom. It is a symmetric distribution. A confidence interval can now be constructed for the population mean M from the selected random sample, say with a confidence coefficient of $(1 - \alpha)\%$. The values of '$t_\alpha$' for different values of $\alpha = Pr.[t > t_\alpha] + Pr.[(- t) < (- t_\alpha)] = 2\ Pr.[t > t_\alpha]$ and different degrees of freedom have been tabulated in, for instance, Rao, C.R. and Others (1966). The confidence interval with a confidence coefficient $(1 - \alpha)$ for the population mean M would be as in *2.19* below – easily computed from the sample observations.

$$[m - t_\alpha\, v(m) < M < m + t_\alpha\, v(m)] \quad (2.20)$$

***We note that the rule 2.16 applies here also except that we use (i) the square root of the unbiased estimate of the sampling variance of the estimate of the population mean in the place of the standard error of the estimate of the population mean, and (ii) the relevant value of the 't' distribution instead of the normal distribution ……*** (2.21)

We have so far dealt with parent populations that are normally distributed. What will be the nature of the sampling distribution of the sample mean when the ***parent*** *population is not normally distributed?* We examine this question in the next sub-section C.

### C) Assessment of Precision–Parent Population has a Non-Normal Distribution

*The Central Limit Theorem* ensures that, even if the population distribution is not normal,

- the sampling distribution of the sample mean will have a mean equal to the population mean *regardless* of the sample size, and
- as the sample size increases, the sampling distribution of the sample mean approaches the normal distribution.

Thus for large 'n' (sample size), say 30 or more, we can proceed with the steps mentioned in sub-section A above. Further, the Student's t-distribution also approaches the normal distribution as 'n' becomes large so that we can use the statistic 't' in sub-section B *as a normally distributed variable with mean 0 and unit variance for samples of size 30 or more.* We may then adopt the procedure outlined in sub-section A.

### D) Determination of Sample Size

Random sampling methods also help in determining the sample size that is required to attain a desired level of precision. This is possible because the standard error and the coefficient of variation C.V. of the estimate, say, sample mean 'm', are functions of 'n', the sample size. C.V. is usually very stable over the years and its value available from past data can be used for determining the

sample size. We can specify the value of C.V. of the sample mean that we desire as, say, C(m) and calculate the sample size with the help of prior knowledge of the population C.V., namely, C. That is,

$$C(m) = C/\sqrt{n}\ ;\ \text{so that } \sqrt{n} = C/C(m),\ \text{or, } n = [C/C(m)]^2 \qquad (2.22)$$

Or we can define *the desired precision in terms of the error that we can tolerate in our estimate* of 'M' (***permissible error***) and link it with the desired value of C(m). Then,

$$n = [2.576C / e]^2\ ,\ \text{where the permissible error } e = |(m - M)| / M. \qquad (2.23)$$

***If the sanctioned budget is F for the survey:*** Let the cost function be of the form $F_0 + F_1 n$, consisting of two components – overhead cost and cost per unit to be surveyed. As this is fixed as F, $F = F_0 + F_1 n$, and the sample size becomes $\mathbf{n = (F - F_0)/F_1}$. The coefficient of variation of C(m) is not at our choice in this situation since it gets fixed once 'n' is determined. We can, however, determine the error in the estimate of 'm' from this sample (*in terms of the RSE of m*), if the population CV, C is known. If further we suppose that the loss in terms of money is proportional to the value of RSE of m, say, Rs. '*l*' per 1% of RSE of 'm', the total cost of the survey becomes, $\mathbf{L(n) = F_0 + F_1 n + l\,(C/\sqrt{n})}$. ***We can then determine the sample size that minimises this new cost (which includes the cost arising out of loss)***. Differentiating L(n) w.r.t n and equating to zero and simplifying,

$$n = [(l/2)\,(C/F_1)]^{2/3} \qquad (2.24)$$

See also the sub-section below on stratified sampling.

### 7.5.4 Methods of Random Sampling

We have so far dealt with random samples drawn form a population. We did not specify the size of the population. We had assumed that the population is infinite in size. In practice, a population may have a size N, however, large. Let us, therefore, consider drawing random samples of size 'n' from a population of size 'N'. We shall consider the following methods of random sampling:

a) Simple Random Sampling (With Replacement) [SRSWR],

b) Simple Random Sampling (Without Replacement) [SRSWOR]

c) Interpenetrating Sub-Samples (I-PSS),

d) Systematic Sampling (sys),

e) Sampling with Probability Proportional to Size (pps)

f) Stratified Sampling (sts),

g) Cluster Sampling (cs) and

h) Multi-Stage Sampling (mss)

We shall indicate in the following sections a description of the above methods, the relevant operational procedure for drawing a sample and the expressions/ formulae for (a) the estimator of the population mean/total/ proportion, (b) the sampling variance of the sample mean/total/population and (c) unbiased estimate of the sampling variance.

**Check Your Progress 2**

1) Briefly describe the meaning of the terms 'sampling distribution of a statistic' and also what is meant by unbiased estimator of a parameter.

……

……

……

2) How is random sampling procedure helpful in correcting for the bias of an estimate? Illustrate this with the help of an example.

……

……

……

3) Explain the terms 'coefficient of variation' and 'relative standard error'.

……

……

……

### 7.5.4.1 Simple Random Sampling with Replacement (SRSWR)

**The method:** This method of drawing samples at random ensures that (i) each item in the population has an equal chance of being included in the sample and (ii) each possible sample has an equal probability of getting selected. Let us select a sample of 'n' units from a population of 'N' units by simple random sampling with replacement (SRSWR). We select the first unit at random, note its identity particulars for collection of data and place it back in the population. We choose at random another unit – this could turn out to be the same unit selected earlier or a different one, note its identity particulars and place it back. We repeat this process 'n' times to get an SRSWR sample of size 'n'. In such a sample one or more units *may* occur more than once. A sample of 'n' distinct units is also possible. It can be shown that the number of possible samples that can be selected by SRSWR method is $N^n$ and that the probability of any one sample being chosen is $1/N^n$.

**Operational procedure for selection of the sample by SRSWR method:** Tables of Random Numbers are used for drawing random samples. These tables contain a series of four-digit (or five-digit or ten-digit) random numbers. Supposing a sample of 30 units is to be selected out of a population of 3000 units. First allot one number from the set of numbers to 0001 to 3000 as the identification number to each one of the population units. The problem of drawing the sample of size 30 then reduces to that of selecting 30 random numbers, one after another, from the random number tables. Turn to a page of the Tables *at random* and start noting down, from the first left-most column of (four or five or ten-digit) random numbers, the first four digits of the numbers from the top of the column downwards. Continue this operation on to the second column till the required sample size of 30 is selected. If any of the random numbers that comes up is more than 3000, reject it. If some numbers (< 3000) get repeated in the process, it means that the corresponding units of the population would be selected more than once, this being sampling with replacement.

**Estimators from SRSWR samples** (using notations set down earlier):

$m_{srswr} = (1/n) \sum_i y_i$ ( $\sum_i$ , i = 1 to n) is an unbiased estimator for M (2.25)

**Note:** If a unit gets selected in the sample more than once, the corresponding value of $y_i$ will also have to be repeated as many times in the summation for calculating $m_{srswr}$ .

Sampling Variance of $m_{srswr}$ : $V(m_{srswr}) = \sigma^2/n = [1/n]\ [E(y_i^{\ 2}) - M^2]$ (2.26)

Standard Error of $m_{srswr}$ : $SE(m_{srswr}) = \sigma / \sqrt{n}$ (2.27)

CV or RSE of $m_{srswr}$ : $C(m_{srswr}) = (1/\sqrt{n})(\sigma / M) = C(y)/\sqrt{n}$. (2.28)

Note that the sampling variance, SE and CV(RSE) of the sample mean in SRSWR is much less than SE and CV of the variable y and these decrease as the sample size increases. The precision of the sample mean in SRSWR, as an estimator of M increases as the sample size increases. However, the extent of decrease in the standard error will not be commensurate with the size of the increase in the sample size. We would need an unbiased estimator of $\sigma^2$, as $\sigma^2$ may not be known. This is

$v(y) = [1/(n-1)][ss]^2$ (2.29);

Therefore, $v(m_{srswr}) = v(y)/n$ (2.30)

an unbiased estimate of Y, or, $Y^*_{srswr} = Nm$ (2.31)

$V(Y^*_{srswr}) = N^2 (\sigma^2 /n)$ (2.32)

$v(Y^*_{srswr}) = N^2 (1/n)[1/(n-1)] \quad (y_i - m_{srswr})^2$ (2.33)

the sample proportion '$p_{srswr}$' is an unbiased estimate of P (2.34)

$V(p_{srswr})$ is PQ/n (2.35); and $v(p_{srswr}) = npq/(n-1)$ (2.36)

$C(p_{srswr})$ -( $= \sqrt{[(1/n)(PQ)]}/ P = [1/\sqrt{n}]\sqrt{[Q/P]}$. (2.37)

Confidence intervals for the population mean/proportion and the sample size for a given level of precision and/or permissible error can now be derived easily.

### 7.5.4.2 Simple Random Sampling without Replacement (SRSWOR)

**The Method:** This method of sampling is the same as SRSWR but for one difference. If a unit is selected, it is ***not*** placed back before the next one is selected. This means that no unit gets repeated in a sample. Operationally, we draw random numbers between 1 and N and if a random number comes up again, it is rejected and another random number is selected. This process is repeated till 'n' *distinct* units are selected. It can be shown that the number of samples of size n that may be selected from a population of 'N' units by this method is ${}_NC_n = N\ !\ /[\ (N-n)\ !\ n\ !\ ] = [N(N-1)\ (N-2)\ \ldots(N-n+1)]\ /\ [n(n-1)\ (n-2)\ldots\ldots 1]$. *The probability $P_{srswor}(S)$ of any one of the samples being chosen is, $1/{}_NC_n$ .*

**Estimators from SRSWOR samples:**

$m_{srswor} = (1/n) \sum_i y_i$ , $\sum_i$ , i =1 to n, is an unbiased estimator of M (2.38)

$V(m_{srswor}) = [(N-n)/(N-1)]\ [\sigma^2/n]$

$= [(N-n)/(N-1)][1/n][(1/N) \sum_i (Y_i - M)^2]$, $\sum_i$, i=1 to N (2.39)

$V(m)_{srswor} < V(m_{srswr})$ since $(N - n) / (N - 1)$ is less than 1 for $n > 1$, (2.40)

Both $m_{srswor}$ and $m_{srswr}$ are unbiased estimators of M *but $m_{srswor}$ is a more efficient estimator of M than $m_{srswr}$* . The factor $[(N - n)/(N - 1)]$ in *(2.40)* is called the ***finite population correction or finite population multiplier.*** *The finite population correction required for finite population need not, however, be used when the sampling fraction (n / N) is less than 0.05.*

$v(m_{srswor}) = [(N - n)/N][1/n][1/(n - 1)][\ _i (y_i - m)^2]$ , $\ _i$ , *i= 1 to n*

$=[(N- n)/N][1/n][1/(n - 1)][ss]^2$, (2.41)

*Unbiased estimate of population total Y, that is* $Y^*_{srswor} = Nm_{srswor}$ (2.42)

$V(Y^*_{srswor}) = N^2\ V(m_{srswor})$ (2.43)

*unbiased estimate of* $V(Y^*_{srswor})$*, namely,* $v(Y^*_{srswor}) = N^2\ v(m_{srswor}$ (2.44)

$C(Y^*_{srswor}) = C(m_{srswor})$ (2.45)

the sample proportion 'p' is an unbiased estimate of 'P' in SRSWOR also. (2.46)

$V(p) = [(N - n)/ (N - 1)]\ [PQ/n]$ , where $P + Q = 1$ (2.47)

$v(p) = [(N - n)/N]\ [pq/(n - 1)]$ , where $p + q = 1$ (2.48)

$C(p) = \sqrt{[(N - n)/(N - 1)]}\ \sqrt{[(1/n)}\ \sqrt{[Q/P]}$ and (2.49)

**Check Your Progress 3**

1) A population has 80 units. The relevant variable has a population mean of 8.2 and a variance of 4.41. These SRSWR samples of size (i) 16, (ii) 25 and (iii) 49 are drawn from the population. What is the standard error (SE) of the sample means from the three samples? Is the extent of reduction in SE commensurate with that of the increase in sample size?

   ……

   ……

   ……

2) What are the results when the sampling method in drawing the three samples in problem 1 above is changed to SRSWOR? What is your advice regarding the choice between increasing the sample size and changing the sampling method from SRSWR to SRSWOR?

   ……

   ……

   ……

3) Indicate whether the following statements are true (T) or false (F). If false, what is the correct position?

   1) The standard error of the sample mean decreases in direct proportion to the sample size. (T/F)

   2) SRSWOR method of sampling is more advantageous than SRSWR for a sampling fraction of 0.02 (T/F)

   3) If Y* = Nm and Variance of m is V(m), the variance of Y* is NV(m). (T/F)

### 7.5.4.3 Interpenetrating Sub-Samples (I-PSS)

Suppose a sample is selected in the form of two or more sub-samples drawn according to the same sampling method so that each such sub-sample provides a valid estimate of the population parameter. The sub-samples drawn in this way are called ***interpenetrating sub-samples (I-PSS).*** This is operationally convenient, as the different sub-samples could be allotted to different investigators. The sub-samples need not be independently selected. There is, however, an important advantage in selecting *independent interpenetrating sub-samples. It is then possible to easily arrive at an unbiased estimate of the variance of the estimator even in cases where the sampling method/design is complex and the formula for the variance of the estimator is complicated.*

Let $\{t_i\}$, i = 1,2,…,h be unbiased estimates of a parameter θ based on 'h' independent interpenetrating sub-samples. Then,

$t = (1/h) \sum_i t_i$ , ( $\sum_i$ , i = 1 to h) is an unbiased estimate of θ (2.50)

$v(t) = [1/h(h-1)] [\sum_i (t_i - t)^2]$, ( $\sum_i$ , i = 1 to h) is an unbiased estimate of V(t) (2.51)

If the unbiased estimator 't' of the parameter θ is *symmetrically distributed* (for example, normally distributed), the probability of the parameter θ lying between the maximum and the minimum of the 'h' estimates of θ obtained from the 'h' sub-samples is given by:

Prob.[Min of $\{t_1, t_2, \text{---- } t_h\} < \theta <$ Max of $\{t_1, t_2, \text{-----} t_h\}] = [1-(1/2)^{(h-1)}]$ (2.52)

This is a confidence interval for θ from the sample. The probability increases rapidly with the number of I-P sub-samples – from 0.5 (two sub-samples) to 0.875 (four sub-samples).

### 7.5.4.4 Systematic Sampling

**The Method:** Let $\{U_i\}$, i = 1,2,……. N be the units in a population. Let 'n' be the size of the sample to be selected. Let 'k' be the integer nearest to N/n - denoted usually as [N/n] — the reciprocal of the sampling fraction. Let us choose a random number from 1 to k, say, 'r'. We then choose the $r^{th}$ unit, that is, $U_r$ . Thereafter, we select every $k^{th}$ unit. In other words, we select the units $U_r$, $U_{r+k}$ , $U_{r+2k}$ ,…… This method of sampling is called *systematic sampling with a random start***.** 'r' is known as *the random start* and 'k' *the sampling interval***.** There would thus be 'k' possible systematic samples, each corresponding to one random start from 1 to k. The sample corresponding to the random start 'r' will be

$$\{U_{r+jk}\}, \quad j = 0,1,2,\ r+jk \le N.$$

The sample size of all the 'k' systematic samples will be 'n' if N = nk. All the 'k' systematic samples will not have a sample size 'n' if $N \neq nk$. For example, if we have a sample of 100 units and we wish to select systematic samples of size 14, the sampling interval is k = [100/15] or 7 The samples with the random starting 1 and 2 will be of size 15 while the other 5 systematic samples (with random starts 3 to 7) will be of size 14.

In systematic sampling, units of a population could thus be selected at a uniform interval that is measured in time, order or space. We can for instance choose a sample of nails produced by a machine for five minutes at the interval

of every two hours to test whether the machine is turning out nails as per the desired specifications. Or, we could arrange the income tax returns relating to an area in the order of increasing gross income returned and select every fiftieth tax return for a detailed examination of the income of assesses of the area. Systematic samples are thus operationally easier to draw than SRSWR or SRSWOR samples. Only one random number needs to be chosen for selecting a systematic sample.

**Estimators from Systematic Samples:**

**An unbiased estimator of the population mean M** based on a systematic sample is given by a **slight variant of the sample mean**, namely,

$m_{sys}^* = (k/N) \quad _i y_i , \quad _i$ , i =1 to n*, n* is the size of the selected sample and k the sampling interval v(2.53)

If N = nk, $m_{sys}^* = m$ the sample mean. If N ≠ nk, there is a bias in using the sample mean as the estimator for M, and

*the bias in using the sample mean as an estimator of M is likely to be small in the case of systematic samples selected from a large population.* (2.54)

The disadvantages, referred to above, in systematic sampling, namely, N not being a multiplier of the sample size n and the sample mean not being an unbiased estimator of the population mean can be overcome by adopting a procedure called ***Circular Systematic Sampling (CSS).*** If 'r' is the random start, and k the integer nearest to N / n, we choose the units.

**$\{U_{r+jk}\}$, if r+jk ≤ N and $\{U_{r+jk-N}\}$, if r+jk > N ; j = 0,1,2,…… ( n - 1).**

Taking the earlier example of selecting a systematic sample of size 14 from a population of 100 units (N == 100, k = 7 and n = 15) all the samples can be made to have a size of 15 by adopting the CSS. A random start of 5 will lead to the selection of a sample of the 15 units 5,12,19,26,33,40,47,54,61, 68.75,82,89,96 and 3 (96 + 7 – 100). This procedure ensures equal probability of selection to every unit in the population.

***Besides constancy of the sample size from sample to sample, the CSS procedure ensures that $m_r$ the sample mean is an unbiased estimate of the population mean.*** (2.55)

Let nk = N. Then m* = m. There are k possible samples, each sample with a probability of 1/k. Let the sample mean of the r-th systematic sample be $m_r = (1/n) \quad _i y_{ir}$, where $y_{ir}$ is the value of the characteristic under study for the i-th unit in the r-th systematic sample, summation is from i = 1 to n. As already noted $m_i$ is an unbiased estimator of M or $E(m_r) = M$. We thus have k possible unbiased estimates of M. Denoting the sample mean in systematic sampling as $m_{sys}$, the sampling variance of $m_{sys}$, and related results of interest are:

$V(m_{sys}) = \sigma_b^2$ (the *between*-sample variance). (2.56)

$V(m_{sys}) = V(y) - \sigma_w^2$ , where $\sigma_w^2$ is *within*-sample variance. (2.57)

Equation *2.57* shows that (i) $V(m_{sys})$ is less than the variance of the variable under study or the population variance, since $\sigma_w^2$ is > 0 and (ii) $V(m_{sys})$ can be reduced by increasing $\sigma_w^2$**,** or by increasing the *within*-sample variance. (ii) would happen ***if the units within each systematic sample are as heterogeneous as***

***possible***. Since we select a sample of 'n' units from the population of N units by selecting every k-th element from the random start 'r', the population is divided into 'n' groups and we select one unit from each of these 'n' groups of population units. Units *within* a sample would be heterogeneous if there is heterogeneity *between* the 'n' groups. This would imply that units *within* each of the n groups would have to be as homogeneous as possible. All these suggest that the sampling variance of the sample mean is related to the *arrangement of the units in the population.* This is both an advantage and disadvantage of systematic sampling. An arrangement that conforms to the conditions mentioned above would lead to a smaller sampling variance or an efficient estimate of the population mean while a 'bad' arrangement would lead estimates that are not as efficient.

### 7.5.4.5 Sampling with Probability Proportional to Size (PPS)

**The Sampling Method:** We have so far considered sampling methods in which the probability of each unit in the population getting selected in the sample was equal. There are also methods of sampling in which the probability of any unit in the population getting included in the sample varies from unit to unit. One such method is sampling with probability proportional to size (pps) in which the probability of selection of a unit is proportional to a given measure of its size. This measure may be a characteristic related to the variable under study. One example may be the employment size of a factory in the past year and the variable under study may be the current year's output. Does this method lead to a bias in our results, as units with smaller sizes would be under represented in the sample and those with larger sizes would be over represented. It is true that if the sample mean 'm' were to be used to estimate the population mean M, m would be a biased estimator of M. However, ***what is done in this method of sampling is to weight the sample observations with suitable weights at the estimation stage to obtain unbiased estimates of population parameters, the weights being the probabilities of selection of the units.***

**Estimates from pps sample of size 1:** Let the population units be $\{U_1, U_2, \ldots\ldots U_N\}$.Let the main variable Y and the related size variable X associated with these units be$\{Y_1, X_1; Y_2, X_2; \ldots\ldots\ldots\ldots Y_N, X_N\}$. The probability of selecting any unit, say, $U_i$ in the sample will be $P_i = (X_i / X)$, where $\sum_i X_i = X$, where $\sum_i$ , i = 1 to N. Let us select *one unit* by pps method. Let the unit selected thus have the values $y_1$ and $x_1$ for the variables y and x. The variables y and x are random variables assuming values $Y_i$ and $X_i$ respectively with probabilities $P_i$ , i = 1,2,……N. The following results based on the sample of size 1 can be derived easily:

An unbiased estimator of population total Y is $Y^*_{(1)\,pps} = y_1 / p_1$ (2.58)

An unbiased estimator of M is $m^*_{(1)pps} = (1/N)\, Y^*_{(1)pps} = (1/N)(y_1 / p_1)$ (2.59)

$V[Y^*_{(1)\,pps}] = \sum_i (Y_i^2 / P_i) - Y^2$ (2.60)

$V[m^*_{(1)pps}] = (1/N^2)\, V[Y^*_{(1)pps}] = (1/N^2) \left[\sum_i (Y_i^2 / P_i) - Y^2\right]$ (2.61)

These show that *the variance of the estimate will be small if the $P_i$ are proportional to $Y_i$.*

**Estimators from pps sample of size > 1 [pps with replacement (pps-wr)]**

A sample of n ( > 1) units with pps can be drawn with or without replacement. Let us consider a pps-wr sample. Let $\{y_i , p_i\}$ be respectively the sample observation on the selected unit and the initial probability of selection at the i–th draw, i = 1,2,----n. Each $(y_i / p_i)$ , i = 1,2,----- n in the sample is an unbiased estimate $[Y_{i(pps\text{-}wr)}^*]$ of the population total Y and $V(Y_{i(pps\text{-}wr)}^*) = \sum_i (Y_r^2 /P_r) – Y^2$, $\sum_r$ , r = 1 to n. (see ***2.60***). Estimates from pps-wr samples are:

$$Y^*_{pps\text{-}wr} = (1/n) \sum_i (y_i / p_i) = (1/n) \sum_i Y^*_{(i)pps\text{-}wr} ; \sum_i , i = 1 \text{ to } n. \quad (2.62)$$

$$V(Y^*_{pps\text{-}wr}) = (1/n)[\sum_r (Y_r^2 /P_r) – Y^2] ; \sum_r , r = 1 \text{ to } N. \quad (2.63)$$

$$V(m_{pps\text{-}wr}) = (1/N^2)(1/n)[\sum_r (Y_r^2 /P_r) – Y^2] ; \sum_r , r = 1 \text{ to } N. \quad (2.64)$$

$$v(Y^*_{pps\text{-}wr}) = [1/\{n(n – 1)\}][\sum_r (y_r^2 / p_r^2 – nY^{*2}]; \sum_r , r = 1 \text{ to } n; \text{ (using 2.51)} \quad (2.65)$$

**Operational procedure for drawing a pps-wr sample:** The steps are:

1) Cumulate the sizes of the units to arrive at the cumulative totals of the unit sizes. Thus

   $T_{i-1} = X_1 + X_2 + \text{-----+ } X_{i-1}$ ; $T_i = X_1 + X_2 + \text{---+ } X_{i-1} + X_i = T_{i-1} + X_i$ ; i = 1, 2,……..., N.

2) Then choose a random number R between 1 and $T_N = X_1 + X_2 +\text{……..+ } X_N = X$ .

3) Choose the unit $U_i$ if R lies between $T_{i-1}$ and $T_i$, that is, if $T_{i-1} < R \leq T_i$ . The probability $P(U_i)$ of selecting the i-th unit will thus be $P(U_i) = (T_i - T_{i-1}) /T_N = X_i / X = P_i$

4) Repeat the operation 'n' times for selecting a sample of size n with pps-wr.

### 7.5.5.6 Stratified Sampling

**The Method:** We might sometimes find it useful to classify the universe into a number of groups and treat each of these groups as a separate universe for purposes of sampling. Each of these groups is called a ***stratum*** and the process of grouping ***stratification***. Estimates obtained from each stratum can then be combined to arrive at estimates for the entire universe. This method is very useful as (i) it gives estimates not only for the whole universe but also for the sub-universes and (ii) it affords the choice of different sampling methods for different strata as appropriate. It is particularly useful when a survey organistion has regional field offices. This method is called ***Stratified Sampling.***

Let us divide the population (universe) of N units into k strata. Let $N_s$ be the number of units in the s-th stratum. $Y_{si}$ be the value of the i-th unit in the s-th stratum. Let the population mean of the s-th stratum be $M_s$. $M_s = (1/N_s) \sum_i Y_{si}$, $\sum_i$, i = 1,2,----,$N_s$ (that is over the units within the s-th stratum) and the population M is $=(1/N) \sum_s N_s M_s = \sum_s W_s M_s$ , where $W_s = (N_s / N)$ and $\sum_s$ being over the strata s = 1,2,……,k). Suppose that we select random samples from each stratum and the sampling method for different strata are different. Let the unbiased estimate of the population mean $M_s$ of the s-th stratum be $m_s$. Denoting 'st' for stratified sampling, **an unbiased estimator of M is given by**

$$m_{st} = \sum_s W_s m_s = (1/N) \sum_s N_s m_s, \quad \sum_s , s = 1 \text{ to } k. \quad (2.66)$$

$$V(m_{st}) = \sum_s W_s^2 V(m_s) = (1/N^2) \sum_s N_s^2 V(m_s), \quad \sum_s , s= 1 \text{ to } k \quad (2.67)$$

$Cov.(m_s,m_r) = 0$ for $s \neq r$ ; (samples from diff. strata are independently chosen) .. (2.68)

$Y_{st}^* = \sum_s Y_s^*$ ; $\sum_s$ , s = 1 to k. (2.69)

$V(Y_{st}^*) = \sum_s V(Y_s^*)$, $\sum_s$, s = 1 to k. (2.70)

Thus estimators with smaller variance (efficient estimators) can be obtained in stratified sampling if we form the strata in such a way as to minimise *intra-strata* or within-strata variation, that is, variance within strata. This would mean maximising between-strata or *inter-strata* variation, since the total variation is made up of within-strata and between-strata variation. In other words, *units in a stratum should be homogeneous.*

Stratified sampling enables us to choose the sample we wish to select by drawing independent samples from each of the different strata in to which we have grouped the universe. *How do we allocate the total sample size 'n' among the different strata?* One way is to *allocate the sample size to different strata in proportion to the size of individual strata measured by the number of units in these strata, namely,* $N_s$, [ $\sum_s N_s = N$, (s= 1 to k)]. This method is especially appropriate in situations where no information is available except the sizes of the strata. The sample sizes for the samples from the stratum, say, the s-th stratum, would then be $n_s = n(N_s/N)$ and $\sum_s n_s$ can easily seen to be equal to 'n'. There are other methods like allocation of the sample size among strata in proportion to the stratum totals of the variable under study, that is, $Y_s$ , the stratum total of the s-th stratum, 's' = 1 to k. We shall not go into the details of other methods here except one situation, namely, ***when we have a fixed budget F sanctioned for the survey.*** Let the cost function F be of the form $F_0 + \sum_s n_s F_s$ , ( $\sum_s$ , s = 1 to k), where $F_0$ , $n_s$ and $F_s$ are respectively the overhead cost, the sample size in stratum 's' and the per unit cost of surveying a unit in stratum 's' (s = 1,2,-----, k). We can determine the optimum stratum-wise-sample-size by minimising the sampling variance of the sample mean (***2.67***) subject to the constraint that the cost of the survey is fixed. The stratum-wise optimum sample size is given by

$$n_s = [(F - F_0)]\ [W_s\sqrt{(V_s / F_s)}] / [\sum_s W_s\sqrt{(V_s F_s)}] \quad , s = 1, k. \qquad (2.71)$$

*The stratum sample size should, therefore, be proportional to* $W_s\sqrt{(V_s / F_s)}$. The minimum variance with the $n_s$ so determined is,

$$\text{Min. } V(m_{st}) = [\ W_s\sqrt{(V_s F_s)}]^2 / (F - F_0)] \qquad (2.72)$$

**Check Your Progress 4**

1) When is PPS method adopted?

……

……

……

2) When will the sampling variance of $Y_{pps}$ * be small?

……

……

……

3) Say True (T) and False (F):

   a) PPS and stratified sampling can be combined with other sampling methods. (T/F)

   b) $V(m_{st})$ is reduced by ensuring that units within individual strataare heterogeneous. (T/F)

   c) The size of a stratified sample can be allocated among the strate, the size being the number of population units in a stratum. (T/F)

4) A systematic sample of size 18 has to be selected from a population of 124. What problems do you face in selecting the sample? Is the sample mean the unbiased estimator of the population mean M? How do you overcome these problems?

   ……

   ……

   ……

### 7.5.4.7 Cluster Sampling

**The method:** Supposing we are interested in studying certain characteristics of individuals in an area. We would naturally select a random sample of individuals from all the individuals residing in the area and collect the required information from the selected individuals. We might also think of selecting a sample of households out of all the households in the area and collect the required details from all the individuals in the selected households. *The households in the area are **clusters** of individuals and what we have done is to select **a sample of such clusters** and to collect the information needed from **all the individuals in the selected clusters** instead of selecting a random sample of individuals from all persons in the area.* What we have done is ***cluster sampling***. ***Cluster Sampling is a process of forming suitable clusters of units and surveying all the units in a sample of clusters selected according to an appropriate sampling method. The clusters of units are formed by grouping neighbouring units or units that can be conveniently surveyed together***. Sampling methods like srswr, srswor, systematic sampling, pps and stratified sampling discussed earlier can be applied to sampling of clusters by treating clusters themselves as sampling units. The clusters can all be of equal size or varying sizes, that is, the number of units can be the same, or vary from cluster to cluster. Clusters can be mutually exclusive, that is, a unit belonging to one cluster will not belong to any other cluster. They could also be overlapping.

**Estimates from cluster sampling:** Let us consider a population of NK units divided into N *mutually exclusive* clusters of K units each – a case of clusters of equal size. The population mean M and the cluster means are given respectively by $M = (1/N) \sum_s m_s$, $\sum_s$ being over clusters s = 1 to N and $m_s = (1/K) \sum_i Y_{si}$, $\sum_i$ being from i = 1 to K within the s-th cluster. Let us draw a sample of one cluster by srs. ***The cluster mean $m_{c\text{-}srs}$ (the subscript c-srs denotes cluster sampling with srs) is an unbiased estimate of M.*** The sampling variance of the sample cluster mean is

$$\mathbf{V(m_{c\text{-}srs}) = (1/N) \sum_s (m_s - M)^2 = \sigma_b^2 = Variance\ \textit{between}\ clusters;\ \sum_s,\ s = 1\ to\ N;} \quad (2.73)$$

Let us compare $V(m_{c\text{-}srs})$ with the sampling variance of the sample mean *when K units are drawn from NK units by SRSWR method.* How does the "sampling efficiency" of cluster sampling compare with that of SRSWR?. ***The sampling efficiency of cluster sampling compared to that of SRSWR, $E_{c/srswr}$ , is defined as the ratio of the reciprocals of the sampling variances of the unbiased estimators of the population mean obtained from the two sampling methods***. The sampling variances and sampling efficiency are

$$V(m_{srswr}) = (1/K)\ [(1/NK) \quad _s \ _i Y_{si}^2 - M^2] = \sigma^2/K \tag{2.74}$$

$$\sigma^2 = \sigma_w^2 + \sigma_b^2 = \textit{within}\text{-cluster variance} + \textit{between}\text{-cluster variance.} \tag{2.75}$$

$$E_{c/srswr} > 1 \ \text{ if } \ \sigma_w^2 > (K-1)\sigma_b^2 \tag{2.76}$$

Thus, cluster sampling is more efficient than SRSWR if the *within*-cluster variance is larger than (K –1) times the *between*-cluster variance. Is this likely? This is not likely as the *between*-cluster variance will usually be larger than the *within*-cluster variance due to *within*-cluster homogeneity. ***Cluster sampling is in general less efficient than sampling of individual units from the point of view of sampling variance***. Sampling of individuals could provide a better cross section of the population than a sample of clusters since units in a cluster tend to be similar.

### 7.5.4.8 Multi-Stage Sampling

We noted in the sub-section on cluster sampling that random sampling of units directly is more efficient than random sampling of clusters of units. But cluster sampling is operationally convenient. How to get over this dilemma? We may first select a random sample of clusters of units and thereafter select a random sample of individual units from the selected clusters. We are thus selecting a sample of units, but from selected clusters of units. What we are attempting is a ***two-stage sampling***. This can thus be a compromise between the efficiency of direct sampling of units and the relatively less efficient sampling of clusters of units. This type of sampling would be more efficient than cluster sampling but less efficient than direct sampling of individual units. In the sampling procedure now proposed, the clusters of units are ***the first stage units (fsu)*** or ***the primary stage units (psu).*** The individual units constitute ***the second stage units (ssu)*** or ***the ultimate stage units (usu)***.

This procedure of sampling can also be generalised to multi-stage sampling. Take for instance a rural household survey. The fsu's in such a survey may consist of districts, the ssu's may be the *tehsils or taluks* chosen from the districts selected in the first stage, the third stage units could be the villages selected from the *tehsils or taluks* selected in the second stage and the fourth and the ultimate stage units (usu's) may be the households selected from the villages selected in the third stage. Such multi-stage sampling procedures help in utilising such information related to the variable under study as may be available in choosing the sampling method appropriate at different stages of sampling. In a multi-stage sampling, estimates of parameters are built up stage by stage. For instance, in two-stage sampling, estimates of the *sample aggregates relating to the fsu's* are built up from the ssu's using the sampling method adopted for selecting the ssu's. These estimates are then used with the sample probabilities of selection of fsu's to build up estimates of the relevant population parameters.

## 7.6 THE CHOICE OF AN APPROPRIATE SAMPLING METHOD

We have considered a number of random sampling methods in the foregoing sub-sections. A natural question that arises now is – ***which method is to be adopted in a given situation?*** Let us consider this question, although the answer to it lies scattered across the foregoing sub-sections. The choice of a sampling design depends on considerations like *a priori* information available about the population, the precision of the estimates that a sampling design can give, operational convenience and cost considerations.

1) When we do not have any *a priori* information about the nature of the population variable under study, SRSWR and SRSWOR would be appropriate. Both are operationally simple. However, ***SRSWOR is to be preferred,*** since $V(m_{srswor}) < V(m_{srswr})$**.** This advantage holds *only when the sampling fraction is not small, or N and n are not large.*

2) Systematic sampling is operationally even simpler than SRSWR and SRSWOR, but it should not be used for sampling from populations where periodic or cyclic trends/variations exist, though this difficulty can be overcome if the period of the cycle is known. $V(m_{sys})$ can be reduced if the units chosen in the sample are *as heterogeneous as possible.* But this will call for a rearrangement of the population units before sampling.

3) When additional information is available about the variable 'y' under study, say, on a variable (size variable) 'x' related to 'y', the pps method should be preferred. The sampling variance of Y* (or m) gets reduced when the probability of selection of units $P_i = ( X_i /N)$ are proportional to $Y_i$ , that is, the size $X_i$ is proportional to $Y_i$ or the variables x and y are linearly related to each other and the regression line passes through the origin. In such cases pps is more efficient than SRSWR. Further, this method can be utilised along with other sampling methods and their relative efficiencies. pps *is operationally simple.* Pps-wor combines the efficiency of SRSWOR and the efficiency-enhancing capacity of pps. However, most of the procedures of selection available, estimators and their variance for pps-wor are complicated and are not commonly used in practice. This is particularly so in large-scale sample surveys with a small sampling fraction, as in such cases sampling without replacement does *not* result in much gain in efficiency. Hence unless the sample size is small, we should prefer pps-wr.

4) Stratified sampling comes in handy when we wish to get estimates at the level of sub-populations or regions or groups. This method also gives us the freedom to choose different sampling methods/designs in different strata as appropriate to the group (stratum) of the population and the opportunity to utilise available additional information relating to the stratum. The sampling variance of estimators can also be brought down by forming the strata in such a way as to ensure *homogeneity of units within individual strata*. In fact, the stratum sizes can be so chosen as to minimise the variance of estimators, when there is a ceiling on the cost of the survey. ***Stratified sampling with SRS, SRSWOR or pps-wr presents a set of efficient sampling designs***.

5) Sometimes, sampling of groups of individual units than direct sampling of units might be found to be operationally convenient. Supposing it is easier to get a complete frame of clusters of individual units than that of units or,

only such a frame, and not that of the units, is available. (e.g. households are clusters of individuals). In such circumstances, cluster sampling is adopted. ***This is in general less efficient than direct sampling of individual units, as clusters usually consist of homogeneous units. A compromise between operational convenience and efficiency could be made by adopting a two-stage sampling design***, by selecting a sample of individual units (second stage units) from sampled clusters (the first stage units). A multi-stage design would be useful in cases where clusters have to be selected at more than one stage of sampling.

6) Finally, we can use the technique of independent I-PSS in conjunction with the chosen sampling design to get at (i) an unbiased estimate of V(m) for any sampling design or estimator of V(m), however complicated, (ii) a confidence interval for 'M' based only on the I-PSS estimates (when the population distribution is symmetrical) and (iii) a tool for monitoring the quality of work of the field staff and agencies.

7) ***SRSWOR, stratified sampling with SRSWOR and, when available information permits, pps-wr and stratified sampling with pps-wr, turn out to be a set of the more efficient and operationally easy designs to choose from. I-PSS can also be used in these designs where possible and necessary***.

**Check Your Progress 5**

1) We wish to study the wage levels of factory labour. What type of sampling method would you adopt for the study and why if (a) just a list of factories is available with the Chief Inspector of Factories of different State Governments, (b) if the list in (a) above also gives the total number of employees in the individual factories at the end of last year and (c) the list also indicates both the kind of product manufactured in the factory along with the information specified in (b) above.

……

……

……

## 7.7 LET US SUM UP

There are broadly three methods of collecting data. The array of tools used for data collection by such methods has expanded over time with the advent of modern technology. Confining data collection efforts to a sample from the population of interest to the study inevitably leads to questions like the use of random and non-random samples. Judgment sampling, convenience sampling, purposive sampling, quota sampling and snowball sampling all belong to the latter group. The absence of a clear relationship between a non-random sample and the parent universe and the presence of the researcher's bias in the selection of the sample render such samples useless for drawing ***valid*** conclusions about the parent population. But these methods are inexpensive and quick ways of getting a preliminary idea of the universe for use in designing a detailed enquiry and in exploratory research. Random samples, on the other hand, are free from such drawbacks and have properties that help in arriving at valid conclusions about the parent population.

The simplest of the sampling methods – **SRSWR** - ensures equal chance of selection to every unit of the population and yields a sample in which one or more units may occur more than once. '$m_{srswr}$' is an unbiased estimator of M. Its precision as an estimator of M increases as the sample size increases. **SRSWOR** yields a sample of *distinct* units. '$m_{srswor}$' is also unbiased for 'M'. ***SRSWOR is a more efficient than SRSWR as*** $V(m_{srswor}) < V(m_{srswr})$. *But this advantage disappears when the sampling fraction is small (< 0.05).* Both provide an unbiased estimator of V(m). An operationally convenient procedure - ***interpenetrating sub-samples (I-PSS)*** – also provides an unbiased estimator of V(m) for any sampling design and estimator for V(m), however complicated.

**Systematic sampling** is a simple and operationally convenient method used in large-scale surveys that requires only a random start and the sampling interval k = [N/n] for drawing the sample. A slight variant of 'm' is unbiased for 'M'. **C*ircular systematic sampling*** takes care of problems that arise when *N/n is not an integer.*. An unbiased estimate of V(m) is not possible but this problem can be tackled easily. ***Systematic sampling is not recommended when there is a periodic or cyclic variation in the population.*** *This problem too can be overcome if the period of the cycle is known.*

An example of methods where the probability of selection varies from unit to unit is **pps**. The "size" could be the value of a variable related to the study variable. In pps, each $y_i / p_i$ , where $y_i$ is the value of the study variable associated with the selected unit and $p_i$ the probability of selection of the unit, is an unbiased estimate (Y*) of the population total Y and [(1/N) Y*] an unbiased estimator of M. As V(Y*) is small if the probabilities $P_i$ are roughly proportional to $Y_i$ , ***pps sampling is more efficient than SRS if the size variable x is proportional to y, that is, x and y are linearly related and the regression line passes through the origin.*** pps sampling can be done with SRSWR, SRSWOR or systematic sampling. In pps-srswr, $[(1/n) \sum_i (y_i / p_i)]$, ( $\sum_i$ i = 1 to n), is an unbiased estimator of Y. This being the mean of n independent unbiased estimates with the same variance V(Y*), v(Y*) can be derived using the I-PSS technique.

**Stratified Sampling** is used when (i) estimates are needed for subgroups of a universe or (ii) the subgroups could be treated as sub-universes. It gives us the freedom to choose the sampling method as appropriate to each stratum. Estimates of parameters are available for the sub-universes (strata) and these can then be combined over the strata to get estimates for the entire universe. ***SE of estimates based on stratified sampling can be small if we form the strata in such a way as to minimise intra-strata variance. Each stratum should thus consist of homogeneous units, as far as possible. Stratum-wise sample sizes can also so chosen as to minimise the variance of estimators***.

Another operationally convenient sampling method, **cluster sampling**, is to sample groups of units or clusters of units at random and collect data from *all* the units of the selected clusters. For example, the household is a cluster of individuals. SRSWR, SRSWOR, pps or systematic sampling can be used for sampling clusters. ***Cluster sampling is, in general, less efficient than direct sampling of units from the point of view of sampling variance. The question here is one of striking a balance between operational convenience and cost reduction on the one hand and efficiency of the sampling design on the other***.

We could improve the efficiency of cluster sampling by selecting a random sample of units from each of the selected clusters - introduce another stage of

sampling. This is **two-stage sampling**. ***This would be more efficient than cluster sampling but less efficient than direct sampling of units.*** **Multi-stage sampling** can also be done. Such designs are commonly used in large-scale surveys as these facilitate the utilisation of information available and the choice of appropriate sampling designs at different stages.

Thus while non-random sampling methods are useful in exploratory research and preliminary work on planning of enquiries, random sampling techniques lead to *valid* judgments regarding the universe. Among random sampling methods, ***SRSWOR, stratified sampling with SRSWOR and, when available information permits, pps-wr and stratified sampling with pps-wr, turn out to be a set of the more efficient and practically useful designs to choose from. I-PSS can also be used in these designs where possible and necessary***.

## 7.8 SOME USEFUL BOOKS & REFERENCES


**Burgess, R.G.(ed) (1982)** : Field Research: A Sourcebook and Field Manual. (Contemporary Social Research 4), George Allen and Unwin, London.

**Des Raj & Chandok P(1998)** : Sampling Theory, Narosa Publishing House, New Delhi.

**Levin, Richard I. & Rubin, David S. (1991)** : Statistics for Mangement, Fifth Edition, Prentice-Hall of India (Private) Limited, M-97 Connaught Circus, New Delhi – 110001.

**Krishniah, P.R. & Rao, C.R. (Eds.) (1988)** : Handbook of Statistics – Vol. 6 : Sampling, North Holland, Amsterdam.

**Mukhopadhyay, Parimal (1998)** : Theory & Methods of Survey Sampling, Prentice-Hall of India Pvt. Ltd., New Delhi.

**Murthy, M.N. (1967)** : Sampling Theory and Methods, Statistical Publishing Society, 204, Barrackpore Trunk Road, Kolkota-700108.

**Rao, C.R., Mitra, S.K. and Matthai, A. (1966)** : Formulae and Tables for Statistical Work, Statistical Publishing Society, Kolkota – 700108.

**Sampath, S. (2005)** : Sampling Theory & Methods, Second Edition, Narosa Publishing House, New Delhi, Chennai, Mumbai, Kolkata.

**Singh, Kultar (2007)** : Quantitative Social Research Methods, Sage Publications (Pvt.) Limited, New Delhi

**Singleton Jr., Royce A. & Straits, Bruce C. (2005)** : Approaches to Social Research, 4th Edition, Oxford University Press, New York – Oxford.

**Viswanathan, P.K. (2007)** : Business Statistics – An Applied Orientation, Darling Kindersely (India) Pvt. Ltd. – licensees of Pearson Education in South Asia.

# 7.9 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

**Check Your Progress 1**

1) Three methods of data collection are: the census and survey method, the observation method, and the experimental method.

2) A function of population values is a parameter; a function of sample observations is a statistic.

$$M = \frac{1}{N}\sum Y_i,\ m = \frac{1}{N}\sum y_i,\ \sigma^2 = \frac{1}{N}\sum Y_i^2 - M^2,\ \text{and } s^2 = \frac{1}{N}\sum y_i^2 - m^2,$$

where M/m is the mean of the population/sample values are the expressions of population/sample means and population/sample variance.

3) The purpose of drawing/studying a sample is to arrive at inferences about the characteristics of the unknown population. Thus, the sample mean is an estimator of the population mean (of the variable under study). The actual value of the estimator when computed from a sample is the estimate.

4) A representative sample is one that possesses the characteristics of the population in the same proportion as in the population. A conscious or subjective selection of units introduces a bias that erodes the sample's capacity to be a true representative of the characteristics of the population. It is for this reason that a random sampling procedure is preferred. Although non-random sampling is drawn without any regard to the rigors of a random sampling procedure, it gives a rough idea of the unknown features of the population under focus. Implemented judiciously, it can serve the objectives of providing preliminary inputs for planning a detailed enquiry later.

**Check Your Progress 2**

1) Sample observations in a random sample are random variables. A statistic is a function of sample observations and is, therefore, also a random variable. This random variable assumes different values in repeated random samples. These values form a frequency distribution of the statistic. This frequency distribution will, in the limit, tend to a probability distribution that gives the probabilities of the statistic assuming different values. This probability distribution is called the sampling distribution of the statistic. Let a statistic 't' be an estimator of a population parameter, say, T. The estimator 't' is an unbiased estimator of the parameter 'T', if the mean of the sampling distribution of 't' is 'T'. Since the mean of the sampling distribution of 't' can also be denoted by E(t), we can express this condition as E(t) = T.

2) Let 't' be an estimator of 'T'. The sampling distribution of the statistic 't' can be derived *only* if the statistic is based on a random sample selected from the population with a parameter 'T'. The mean of the sampling distribution can then be derived from the sampling distribution of 't'. That is, E(t) can be derived. If E(t) is $\neq$ T and E(t) = KT or T+K, we can correct the estimator 't' as (t/K) or (t – K) to arrive at an unbiased estimator of T. An example is the unbiased estimator of $\sigma^2$.

3) See Sub-section 7.5.3 A

**Check Your Progress 3**

1) The answer for (i) would be $\sqrt{\frac{4.41}{16}} = \frac{2.1}{4} = 0.525$. The answer for (ii) is $\sqrt{\frac{4.41}{25}} = \frac{2.1}{5} = 0.42$. The percentage decrease in SE works out to [(0.525 – 0.42)x 100] / [0.525] = 20%. The percentage increase in the sample size is $= \frac{9}{16} * 100 = 56.25\%$ . The extent of increase is larger than the extent of decrease in the standard error. Try to work out (iii) and the compare the results of (ii) & (iii).

2) Use the formula 2.39 for *variance* of the sample mean under SRSWOR. Note that, *for the same sample size,*

$$V(m_{srswor}) - \frac{N-n}{N-1} V(m_{srswr}) = 1 - \frac{n-1}{N-1} V(m_{srswr}).$$

For n =16 and using SRSWOR sampling, SE works out to $\frac{64}{79} * 0.525 = 0.425$.

Compare it with the SE for n = 25 in problem 1. SE of the sample mean was reduced to more or less this level by *increasing the size of the SRSWR sample from 16 to 25*. Work out (ii) & (iii)

3) a) F. SE(m) decreases in inverse proportion to the square root of the sample size.

b) F. SRSWOR is more advantageous if the sampling fraction is > 0.5.

c) F. $V(Y^*) = N^2 V(m)$.

**Check Your Progress 4**

1) When information on an auxiliary variable ‘x’ related to the variable ‘y’ under study is available for each unit of the population.

2) When $P_i$ are proportional to$Y_i$ . That is, when the variable ‘x’ and ‘y’ are linearly related to each other and the regression line passes through the origin.

3) (a) T; (b) F: they should be homogeneous; (c) T.

4) The integer ‘k’ nearest to 124/18 is 7. Choose a random start between 1 and 7 for selecting the sample. Since N = 124 is not a multiple of n = 18, two problems arise. (i) Samples with random start 1 to 5 are of size 18 and others of size 17. The statistic $\frac{k}{N}\sum y_i = \frac{7}{124}\sum y_i$ and *not* the sample mean is an unbiased estimate of M. Both the problems at (i) and (ii) are overcome by adopting the circular systematic sampling procedure.

**Check Your Progress 5**

1) a) State-wise lists are available. Stratified sampling with each State (or part of it depending on the number of factories) being a stratum is an obvious choice. We need to collect details relating to *workers* in the factories. Within each stratum a cluster of factories can be selected and wage data relating to *all the workers* in each selected factory may be collected. Alternatively, we could adopt a two-stage sampling – select a sample of factories (fsu’s) and a sample of workers (ssu’s) from the selected factories, to enhance the efficiency of sampling. SRSWOR would be appropriate for each stage of selection.

# UNIT 8 MEASUREMENT AND SCALING TECHNIQUES

**Structure**



## 8.0 OBJECTIVES

After going through this Unit, you will be able to:

- state the concept of measurement and its need;
- explain the various scales of measurement;
- discuss the various scaling techniques; and
- identify the criteria for good measurement.

## 8.1 INTRODUCTION

Measurement is the foundation of scientific enquiry. As we know that research begins with a 'problem' or topic. Thinking about a problem results in identifying concepts that capture the phenomenon being examined. Concepts are mental ideas representing the phenomenon. The process of conceptualization involves defining the concepts abstractly in theoretical terms. Operationalisation of concepts for research purpose involves moving from the abstract to the empirical level. This underlines the need to measure the various attributes of the people, the characteristics of objects or phenomenon. Further issues like decent work, human wellbeing, happiness, quality of education etc. which have several qualitative and quantitative dimensions are emerging important issues of research in economics. Such issues are being probed

through development of indicators and composite indexes which necessarily involve measurement of such indicators. Hence, a student of economics is expected to be well versed in the measurement scales, criteria of good measurement, and important scaling techniques. This unit throws light on these issues. Let us begin to discuss the concept of measurement.

## 8.2 CONCEPT OF MEASUREMENT

Simply speaking the process of assigning numbers to various attributes of people, objects or concepts is known as measurement. Technically, measurement is a process of mapping aspects of a domain to other aspects of a range according to some rule of correspondence.**Tyler(1963)** defines measurement as, "assignment of numerals according to rules". **Nunally (1970)**viewed that, "Measurement consists of rules, for assigning numbers to objects in such a way to represent quantities of attributes". Thus, Nunally focuses on both rules and manner in which numbers are assigned to an object.

According to **Campbell**, "Measurement is defined as the assignment of numerals to objects or events according to rules". The fact that numerals can be assigned under different rules leads to different kinds of scales and different kinds of measurement.

Thus, measurement is a process to assign numbers or other symbols to characteristics of objects according to certain rules. Assigning numbers permit statistical analysis of the resulting data and facilitate the communication of measurement rules and results. The rules for assigning numbers hence need to be standardized and uniformly applicable. In the assignment process, there must be one to one correspondence between the numbers and characteristics being measured. In this process, variables can be divided into two basic types – quantitative andqualitative variables.

## 8.3 MEASUREMENT ISSUES IN RESEARCH

A number of concerns arise in the process of measurement which need to be addressed by the researcher. Some of the important issues include the following:

1) Whether the underlying characteristics of the concept allows ordering (ordinal level) or categorizing (nominal level)?
2) Whether the features of the concept are discrete or continuous with fine gradation?

   The answer of these two questions will enable to determine the level of measurement inherent to the concept. As a general rule, efforts should be made that measurement represent the highest scale of measurement for a concept because it allows the use of more powerful statistical techniques for analysis.
3) Another important issue pertains to number of indicators for measurement of a concept. Some simple concepts can be measured by one indicator whereas abstract concepts are measured with more than one indicators. How many indicators are appropriate to measure a concept is to be decided by the researcher.
4) Another measurement issue concerns the source of valid and reliable measure. Hence, considerable attention should be given to identify valid and reliable measures for the study.

5) Measurement should be free from measurement errors like invalidity error and unreliability.

6) Measurement validity is distinct from internal validity and external validity because these two are separate research design issues.

7) A final measurement issue concerns the proper use of available data. For that, a research should be well aware of all available data sources and the types of data available with the various data compilation agencies.

## 8.4 SCALES OF MEASUREMENT

Before discussing different levels of measurement, let us remember that there are three postulates of measurement: (i) Equalities or identities, (ii) Rank order, (iii) Additivity.

You were introduced the elementary concept and types of scales of measurement in Unit 1. Even at the cost of repetition let us recall that Stanley S.Stevens (1946) in his seminal paper, "On the theory of scales of measurement" published in "Science" classified types of scales into four categories: (1) Nominal, (2) Ordinal, (3) Interval, (4) Ratio.

1) **Nominal Scale**

A qualitative scale without order is called nominal scale. In nominal scale numbers are used to name identity or classify persons, objects, groups, gender, industry type. Even if we assign unique numbers to each value, the numbers do not really mean anything. This scale neither has any specific order nor it has any value. In case of nominal measurement, statistical analysis is attempted in terms of counting or frequency, percentage, proportion, mode or coefficient of contingency. Addition, subtraction, multiplication and division are not possible under this scale.

2) **Ordinal scale**

A qualitative scale with order is called ordinal scale. In ordinal scale numbers denote the rank order of the objects or the individuals. Numbers are arranged from highest to lowest or lowest to highest. For example, students may be ranked $I^{st}$, $2^{nd}$ and $3^{rd}$ in terms of their academic achievement. The statistical operations which can be applied in ordinal measurement are median, percentiles and rank correlation coefficient. Ordinal scales do not provide information about the relative strength of ranking. This scale does not convey that the distance between the different rank values is equal. Ordinal scales are not equal interval measurement. Further this scale does not incorporate absolute zero point.

3) **Interval scale**

Interval scale includes all the characteristics of the nominal and ordinal scale of measurement. In interval scale numerically equal distance on the scale indicate equal distances in the attributes of the object being measured. For example, a scale represent marks of students using the attributes range 0 to 10, 10 to 20, 20 to 30, 30 to 40 and 40 to 50, and so forth. The mid point of each range (ie. 5, 15, 25, 35 and 45 etc.) are equidistance from each other. The data obtained from an interval scale is known as interval data. The appropriate measures used in this scale are: arithmetic mean, standard deviation, Karl Pearson's coefficient of correlation and tests like t-test and f-test. We cannot apply coefficient of variation in the interval scale.

4) **Ratio Scale**

Ratio scale has all the characteristics of nominal, ordinal and interval scale. It also possesses a conceptually meaningful zero point in which there is a total absence of the characteristic being measured. Ratio scales are common among physical sciences rather than among social sciences. The examples of ratio scales are the measures of height, weight, distance and so on. All measures of central tendencies that can be used in this scale include geometric and harmonic means.

The properties of these four scales can be summarized in following tabular form:

| Properties / Scale | Category | Ranking | Equal interval | Zero Point |
|---|---|---|---|---|
| Nominal | ✓ | | | |
| Ordinal | ✓ | ✓ | | |
| Interval | ✓ | ✓ | ✓ | |
| Ratio | ✓ | ✓ | ✓ | ✓ |

| Properties / Scale | Indicates Difference | Indicates Direction of Difference | Indicates Amount of Difference | Absolute Zero |
|---|---|---|---|---|
| Nominal | × | | | |
| Ordinal | × | × | | |
| Interval | × | × | × | |
| Ratio | × | × | × | × |

You will notice in the above tables that only the ratio scale meets the criteria for all four properties of scales of measurement. Interval and Ratio scale data are sometimes referred to as parametric and Nominal and Ordinal data are referred to as non-parametric.

**Examples of Scales of Measurement –**

| Scale | Example | Statistics |
|---|---|---|
| Nominal | Gender<br>Yes-no<br>Students roll number<br>Objects/Groups | Frequency, percentage, proportion Mode, Coefficient of contingency, chi-square ($x^2$) |
| Ordinal | Class Rank,<br>Socio Economic status<br>Academic Achievement | Median, percentile, rank, correlation, Range |
| Interval | Student grade point, Temperature, Calendar dates, Rating Scale | Mean, correlation, Range, Standard deviation, rank order variance, Karl pearson's correlation, t-test, f-test etc. |
| Ratio | Weight, Height, Salary, Frequency of buying a product | Mean, Median, Mode, Range, Variance, Standard deviation, coefficient of variation, rank order variance, Karl pearson's correlation, t-test, f-test |

## 8.5 CRITERION FOR GOOD MEASUREMENT

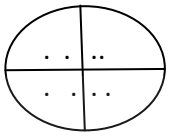
There are three major criteria for evaluating measurement:

1) Reliability
2) Validity
3) Practicality

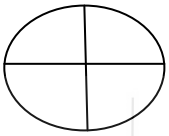
### 8.5.1 Reliability

Reliability refers to the degree to which the measurement or scale is consistent or dependable. If we use same construct again and again for measurement, it would lead to same conclusion. Reliability is consistency in drawing conclusion.

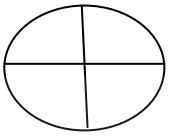
**Reliability and Validity**



Reliable and Valid



Valid but not Reliable



Reliable but not valid

### 8.5.2 Validity

Validity refers to the extent an instrument or scale tests or measures what it intends to measure. This means validity is the extent to which differences found with a measuring instrument reflect true differences among those being tested. This can be done by seeking other relevant evidences. There are three types of validity: content validity; criterion validity and construct validity. Content validity indicates the extent to which it provides coverage of the issues under study. Criterion validity examines how well a given measure relates to one or more external criterion based on empirical observations. And, construct validity explains the variation observed in tests conducted on several individuals. Construct validity is closely related with factor analysis.

### 8.5.3 Practicality

From practical view point, measure should be economical, convenient and interpretable. The measure should not be lengthy and difficult. The measure can be done by highly specialized persons.

## 8.6 ERRORS IN MEASUREMENTS

Measurement need to be precise and unambiguous for validity and reliability of any research study. Hence, measurement should be either free from errors or have minimum error. Any measurement usually involves two types of errors:

- Measurement invalidity
- Unreliability

**Measurement invalidity** refers to the degree to which the measure incorrectly captures the concept.

**Unreliability** refers to inconsistency in what the measure produces under repeated uses. If any measure, on average give some score for a case on variable and gives other score, when it is used again, it is said to be unreliable.

The possible sources of error are:

a) **Respondent** by not expressing strong negative feelings.

b) **Situation** wherein interviewer and respondent are not in good rapport.

c) **Measurer** – errors may also creep because of incorrect coding, faulty tabulations and statistical calculations or behavior or attitude of the interviewer.

d) **Instrument** errors may also arise because of the defective measuring instrument.

**Check Your Progress 1**

1) What do you mean by measurement?

……

……

……

……

……

2) Give examples of interval scale and ratio scale.

……

……

……

……

……

3) How will you examine the validity of a measurement?

……

……

……

……

……

4) Identify three measure concerns that need to be addressed in the process of measurement.

……

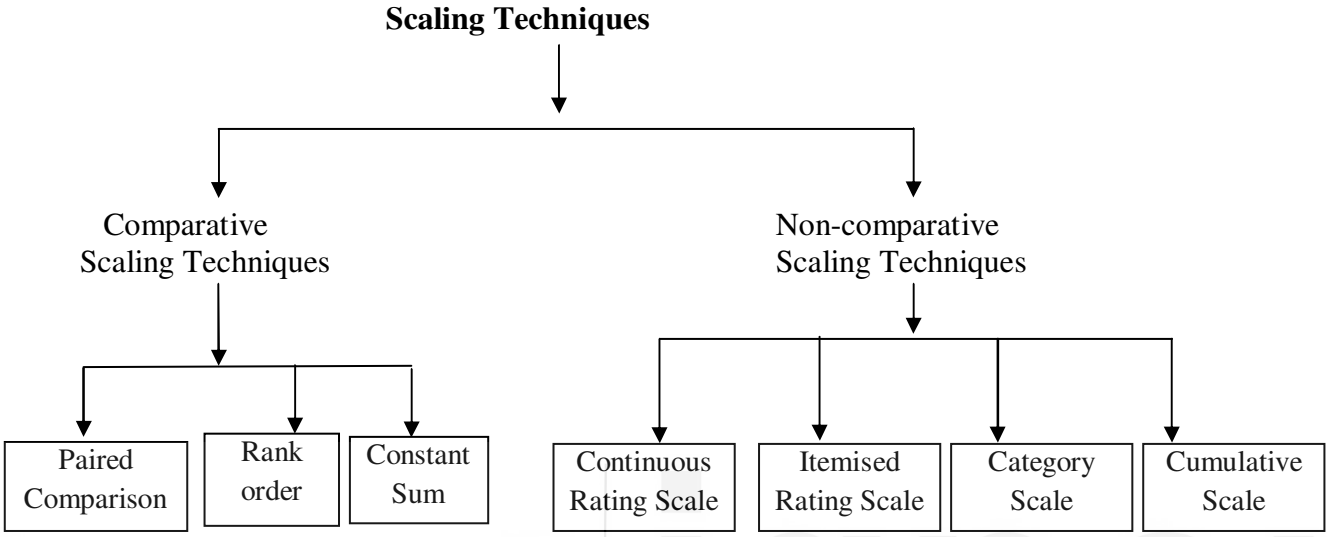
……

……

……

## 8.7 SCALING TECHNIQUES

Scaling describes the way of generation of a continuum upon which measured objects are located. Broadly, scaling techniques can be classified into two categories: comparative, and non comparative scaling.




### 8.7.1 Comparative Scaling Techniques

Comparative scales involve the comparison of objects directly with one another. For example, in a study of consumer preferences for different telecom services, a respondent may be asked to rank choice according to his preferences. In this technique, data must be interpreted in relative terms and have only ordinal or rank order properties. In comparative scales small differences between stimulus objects can be detected.

**Example**

Rank the following factors in order of your importance in choosing mobile services (Assign 1 to most important and 5 to least important factor. Please do not repeat the ranks.)

Factor: Price, connectivity, call drop, Internet, Download speed.

Rank :…… …… …… ……. …….

There are various comparative scaling techniques. Three most commonly used comparative scaling techniques are discussed here.

a) Paired comparison

b) Rank order

c) Constant sum

a) **Paired Comparison**

The paired comparison is most commonly used scaling technique. This method is simply a binary choice. In this method respondents choose the stimulus or items in each pair that has the greater magnitude on the choice dimension they are instructed to use. This method is used when the study requires to distinguish between the two objects. The data obtained in this method is ordinal in nature.

**Example**

In a study of consumer preferences for the two brands of milk product i.e. Amul and Mother Dairy, a consumer is asked to indicate which one of the two brands he would prefer for personal use:

1) Which milk product of the following you prefer on the basis of taste. Please tick mark (√).

| Amul | Mother Dairy |
|---|---|
| | |

2) Which milk product will you prefer on the basis of packaging? Please tick mark (√).

| Amul | Mother Dairy |
|---|---|
| | |

3) Which product will you prefer on the basis of price? Please tick mark (√).

| Amul | Mother Dairy |
|---|---|
| | |

This technique is useful when the researcher wants to compare two or more than two objects. If there are more than two objects (for example n objects) to compare, the total comparison will be;

$$\text{Number of comparison} = \frac{n[n-1]}{2};$$

n= number of objects.

b) **Rank Order**

This method is very popular among researchers and provides ordinal data. In this method, respondents are provided various objects and asked to rank the objects in the list. Rank order method is less time consuming. In this method, if there are n objects, only (n-1) decisions need to be made. Respondent can easily understand the instructions for ranking. This technique produces ordinal data.

**Example**

Rank the various brands of mobile phone in order of preference. The most preferred can be ranked 1, the next as 2 and so on. The least preferred will have the last rank. No two brands should receive the same rank number.

| Sl. No. | Brand | Rank |
|---|---|---|
| 1 | Nokia | |
| 2 | Motorola | |
| 3 | Samsung | |
| 4 | HTC | |
| 5 | Karbonn | |
| 6 | Spice | |
| 7 | Xolo | |
| 8 | Gionee | |
| 9 | LG | |
| 10 | Sony | |

c) **Constant Sum**

In constant sum scaling, respondents are asked to allocate a constant sum of units to a specific set of objects on the basis of pre-defined criterion. If an object is not important, the respondent can allocate zeropoint and if an object is most important may allocate maximum points out of the fixed points. The total fixed points are 100. The total may be taken as some other value depending on the study.

**Example**

Allocate preferences regarding a movie based on various predefined criteria. Respondents were asked to rate each criteria in such a way that the total becomes 100.

| Criteria | Respondent Preference |
|---|---|
| Content Quality | 20 |
| Music | 20 |
| Sound | 15 |
| Story | 30 |
| Breadth of Coverage (local, regional Supplements, national & global) | 15 |
| **Total** | **100** |

The data obtained in this method is considered an ordinal scale. The advantage of this method is that it provides fine discrimination among objects without consuming too much time. However, allocation over large number of objects may create confusion for the respondent. This method cannot be used as a response strategy with illiterate people and children.

## 8.7.2 Non Comparative Scaling Techniques

Non-comparative scaling techniques involve scaling of objects independent to some specific standard. Respondent evaluate only one object at a time. For example, in a study of consumer preferences for different telecom services, a consumer may be asked to rate a list of factors that he/she would consider while choosing a particular telecom company.

**Example**

Please rate the following factors you think important in choosing mobile service. Rate 1 to least important and 5 to most important (rating could be on any scale. In this example, we have used a 5 point scale).

| Factor | Price | Connectivity | Call Drop | Download Speed | Internet |
|---|---|---|---|---|---|
| Rating | | | | | |

In this scaling technique, data is usually in interval scale. It can be continuous, metric/numeric also. Some of the commonly used non-comparative scaling techniques are –

i) Continuous Rating Scale

ii) Itemised Rating Scale

iii) Category Scale

iv) Cumulative Scale

i) **Continuous Rating Scale**

This rating scale is also known as graphic rating scale.In continuous rating scale, respondents indicate their rating by marking at appropriate position on a line. The line is labeled at both ends from one extreme criterion to the other. The line may contain points 0,10,20,…..,100.

**Example**

How would you rate a magazine with regard to its quality.

| Quality Indicators | Scale Measurement |
|---|---|
| Content Coverage | Most_____x_____Least<br>Above 80 60 40 20 0 |
| Language | Most_____x_____Least |
| Presentation Style | Most_____x_____Least |

A very large number of ratings are possible if the respondents are literate to understand and accurately differentiate the objects. The data generated from continuous rating scale can be treated as numeric and interval data. The researcher can divide the line into as many categories as desired and assign scores based on the categories under which the ratings fall.

ii) **Itemized Rating Scale**

In this scale, respondents are provided with a scale having numbers/ descriptions associated with each category. The respondents are asked to select the best fitting category with the object.The commonly used itemised rating scale are:

a) Likert Type Scale (Summated Scale)

b) Semantic Differential Scale

c) Stapel Scale

a) **Likert Type Scale (Summated Scale)**

Likert (1932) proposed a simple and straight forward method for scaling attitudes that is most frequently used today. This scale is also known as summated rating scale. Summated scales consist of a number of statements which express either a favourable or unfavourable attitude towards the given object to which the respondant is asked to respond to each of the statements in terms of several degrees of agreement or disagreement. Let us understand this method with an example.

Against the following statements relating to psychological well-being, please tick mark any one of the following options shown against the statement

| Sl. No. | Statements | Strongly Agree | Agree | Undecided | Disagree | Strongly Disagree |
|---|---|---|---|---|---|---|
| | | 5 | 4 | 3 | 2 | 1 |
| 1 | I always live in present and do not worry about future. | | | | | |
| 2 | Many people think without purpose in life but I am not such type of persons. | | | | | |
| 3 | I generally feel sense of perfection during the moments I live in. | | | | | |
| 4 | I am very well in managing and discharging most of my responsibilities. | | | | | |
| 5 | I give importance to the innovative ideas and experiences that challenge my thought about me and the world. | | | | | |

Depending on the wording of an individual item, an extreme answer of strongly agree or strongly disagree will indicate the most favorable response on the underlying attitude measured by the questionnaire. This scale is relatively easy and quick to compute. The data in this scale is of interval scale.

Some of the important advantages of this scale are: (i) it is relatively easy to construct as it can be performed without a panel of judges, (ii) it is considered more reliable because respondents answer each statement included in the

instrument, (iii) it can easily be used in respondent centred and stimulus centred studies, (iv) this takes less time to construct. However, there are several limitations of this type of scale i.e. (i) with this scale, we can simply examine whether respondents are more or less favourable to a topic but it is difficult to tell how much more or less they are, (ii) This does not rise more than to a ordinal scale structure, (iii) the total score of an individual respondent has little clear meaning since a given total score can be secured by a variety of answer patterns.

b) **Semantic Differential Scale**

This scale was developed by Osgood, suit &Tannenbaum (1957). Semantic differential scale includes a seven point scale in comparison to Likert's five point scale. The semantic differential scale is based on the proposition that an object can have several implied or suggestive meanings to an expressed opinion. The semantic differential scale can be scored on either -3 to +3 or 1 to 7. This scale can also provide interesting comparison between products organizations and so on. The results of this scale are further analysed by the factor analysis.

Example: A researcher has developed an item seeking your perceptions regarding magazine. Please mark√on each line that best indicates your perception. Please be sure to mark every scale and do not omit any scale.

India Today Magazine is

| | +3 +2 +1 0 -1 -2 -3 | |
|---|---|---|
| Easy to read | ____ : ____ : ____:____:_√__:____:____ | Hard to read |
| Modern | ____ : ____ : ____:____:_√__:____:____ | Old fashion |
| Rational | ____ : ____ : ____:____:____:____:____ | Emotional |
| Unreliable | ____ : ____ : ____:____:____:____:____ | Reliable |
| Worthless | ____ : ____ : ____:____:____:____:____ | Valuable |
| Unorganized | ____ : ____ : ____:____:____:____:____ | Organized |
| Orthodox | ____ : ____ : ____:____:____:____:____ | Liberal |
| Vain | ____ : ____ : ____:____:____:____:____ | Modest |

Semantic differential scale provides a very convenient and quick way of gathering impressions on one or more than one concept. The data generated from this scale can be considered as numeric in some cases, and can be summed to arrive total scores adjectives. It must define a single dimension and each pair must be bipolar opposites labeling the extremes.

Like Likert Type Scale, this scale has also several advantages: (i) it is an efficient and easy way to secure attitude from a large sample, (ii) the total set of responses (both direction) gives a comprehensive picture of the meaning of an object as well as a measure of the subject of the rating, (iii) it is a standarised technique that can be easily repeated.

However, this technique escapes many of the problems of response distortion found with more direct methods.

c) **Stapel Scale**

Stapel scales are named after John stapel who developed these scales. Stapel scale consists of a single criterion in the centre with 10 categories numbered from -5 to +5 without a neutral point (zero). The scale is usually presented vertically. Negative rating indicates that the respondent inaccurately describes the object and positive rating indicates that the respondent describes the object accurately.

**Example**

Rate the Departmental store by marking (√) on the following factors. +5 indicate that the factor is most accurate for you and -5 indicate that the factor is most inaccurate for you.

| | | | |
|---|---|---|---|
| +5 | | -5 | |
| +4 | | -4 | |
| +3 | | -3 | |
| +2 | | -2 | |
| +1 | | -1 | |
| **Services** | | **Products** | |
| -1 | | +1 | |
| -2 | | +2 | |
| -3 | | +3 | |
| -4 | | +4 | |
| -5 | | +5 | |

In this scale, data generated is interval data. In this method, data can be collected through telephonic interview. The data obtained from Stapel scale can be analysed in the same way as semantic differential scale.

iii) **Category Scale**

In Category scaling method, objects are grouped into a predetermined number of categories on the basis of their perceived strength along certain dimension. This scale is a dichotomous scale. This method is useful for socio demographic questions. The response is this category we get, typically 'Yes' or 'No' type. Data in this category is either nominal or ordinal. Under this category scale, instructions and response tasks are quick and simple, and many options can be included.

**Example**

A) **Single Category Scale**

1) Do you own a house?

| Yes | No |
|---|---|
| | |

2) Do you own a car?

| Yes | No |
|---|---|
| | |

3) Do you own a Mobile phone?

| Yes | No |
|---|---|
| | |

b) **Multiple Category**

1) Please indicate your income by marking (√) in the income group you fall

| | |
|---|---|
| 5000 - 10000 | |
| 10000 - 15000 | |
| 15000 - 20000 | |
| 20000 - 25000 | |

iv) **Cumulative Scale (Guttman Scale)**

Like other scales, it consists of series of statements to which a respondent expresses his agreement or disagreement. The statements are in a form of cumulative series i.e. in a way, an individual who replies favourably to say item no. 3 also replies favourably to item no. 2 and 1 and so on. The individual's score is worked out by counting the number of points concerning the number of statements he answers favourably. Knowing the total score, we can estimate as to how a respondent has answered individual statements constituting cumulative scales.

Let us understand with the illustration of an **example**.

Here is an example of a Guttman scale - the Bogardus Social Distance Scale:

(Least extreme)

1) Are you willing to permit immigrants to live in your country?
2) Are you willing to permit immigrants to live in your community?
3) Are you willing to permit immigrants to live in your neighbourhood?
4) Are you willing to permit immigrants to live next door to you?
5) Would you permit your child to marry to an immigrant?

(Most extreme)

Cumulative scale (scalogram analysis) like any other scaling technique has several advantages as well as limitations. The advantages are – It assures that only a single dimension of attitude is measured, (ii) Researcher's subjective judgement is not allowed to creep in the development of scale since the scale is determined by the replies of respondents.

Its main limitation is that (i) In practice perfect cumulative or unidimesional scale are very rarely found and approximation is used, (ii) Its developmental procedure is cumbersome in comparison to other scaling methods.

**Check Your Progress 2**

1) What is the distinction between comparative measuring scale and non comparative measuring scale?

   ……

   ……

   ……

   ……

2) What do you mean byLikert Scale? Give its two examples.

   ……

   ……

   ……

   ……

3) What are the sources of measurement errors?

   ……

   ……

   ……

## 8.8 LET US SUM UP

Operationalisation of concepts for research purpose need measurement of various attributes of the people, the characteristics of objects or phenomenon. Measurement is a process to assign numbers to the characteristics according to certain rules. Measurement scales are of four types – nominal, ordinal, interval and ratio. A number of concerns arise in the process of measurement. Important among them are: the scale measurement i.e. the validity, reliability and practicality, number of indicators for measurement etc. Scaling techniques are broadly of two types – comparative, and non-comparative. Within comparative category, paired, rank order and constant sum are important techniques. Within non-comparative techniques, four sub-categories i.e. continuous rating scale, itemized rating scale, category scale and cumulative scale have been covered in this unit. All these scaling techniques have their relative advantages and limitations and are used depending upon the goal and the nature of characteristics of objects/phenomenon to be measured.

## 8.9 KEY WORDS

**Measurement** : The assignment of numerals to objects or events according to rules.

**Nominal** : Numbers are used to name identity or classify persons, objects, groups or gender.

| | | |
|---|---|---|
| **Ordinal Scale** | : | Scale with order is called ordinal scale. |
| **Interval Scale** | : | In this scale numerically equal distance on the scale indicate equal distances in the attributes of the object being measured. |
| **Ratio Scale** | : | A conceptually meaningful zero point in which there is a total absence of the characteristics being measured. |
| **Validity** | : | The extent to which an instrument or scale tests or measures what the measurement or scale i.e. consistent or dependent. |
| **Reliability** | : | The degree to which the measurement or scale i.e. consistent or dependent. |
| **Practicality** | : | Measure should be economical convenient and interpretable. |
| **Errors in Measurement** | : | Discrepancies between the obtained scores and the corresponding scores. |
| **Scaling** | : | The way of generation of a continuum upon which measured objects are located. |
| **Comparative Scaling** | | The comparison of objects directly with one another. |
| **Paired Comparison** | : | A binary choice, respondents choose the objects in each pair that has the greater magnitude on the choice dimension. |
| **Rank Order** | : | Respondents are provided with various objects and asked to rank the list of objects. |
| **Constant Sum** | : | Respondents are asked to allocate a constant sum of units to a specific set of objects with respect to predefined criterion. |
| **Non Comparative Scaling** | : | Scaling of object independently to some specify standard. |
| **Continuous Rating Scale** | | Respondents indicate their rating by marking at appropriate position on a line. |
| **Itemised Rating Scale** | | Respondents are asked to select the best fitting category with the object. |
| **Likert Scale** | : | Respondents are asked to select the best fitting category with the object. |
| **Semantic Differential Scale** | | Seven point scale based on the proposition that an object can have several implied or suggestive meanings to an expressed opinion. |
| **Stapel Scale** | : | A single criterion in the centre with 10 categories numbered from -5 to +5 without a neutral point. |
| **Category Scale** | : | Objects are grouped in to a predetermined number of categories on the basis of their perceived strength along certain dimension. |

## 8.10 SOME USEFUL BOOKS/PAPERS

Stevens S.S. (1946): *On Theory of Scales of Measurement, Science New Series,* Vol. 103, No. 2684, pp 77-680.

Kothari C.R. (2008): *Research Methodology Methods and Techniques,* New Age International Publishers. Chapter 5 page 69 to 82.

Babbie Earl, (2010): *The Practice of Social Research*, 12th Edition, WodsworthCengage Learning, CA, USA.

Malhotra, N.K. and S. Dash (2009): *Marketing Research,* An Applied Orientation, Pearson Education, New Delhi.

Singh, A.K.: *Tests Measurements and Research Method in Behavioural Sciences,* Bharti Bhawan Publishers and Distributors, Patna.

Gregory R.J. (2006): *Psychological Testing History, Principles and Applications,* Pearson Education, New Delhi, India.

Michael S. Lewis-Beck (ed) (2004): *The Sage Encyclopedia of Social Science Research Methods,* Vol. 2, Page no. 161 to 165, Sage Publications, New Delhi, India.

## 8.11 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

**Check Your Progress 1**

1) See Section 8.2
2) See Section 8.4
3) See Section 8.5
4) See Section 8.3

**Check Your Progress 2**

1) See Sub-section 8.7.1 and 8.7.2
2) See Sub-section 8.7.2 under the subhead Itemised Rating Scale
3) See Section 8.6.