Department of Commerce

University of Calcutta

Study Material

Cum

Lecture Notes

Only for the Students of M.Com. (Semester IV)-2020

University of Calcutta

(Internal Circulation)

Dear Students,

Hope you, your parents and other family members are safe and secured. We are going through a world-wide crisis that seriously affects not only the normal life and economy but also the teaching-learning process of our University and our department is not an exception.

As the lock-down is continuing and it is not possible to reach you face to face class room teaching. Keeping in mind the present situation, our esteemed teachers are trying their level best to reach you through providing study material cum lecture notes of different subjects. This material is not an exhaustive one though it is an indicative so that you can understand different topics of different subjects. We believe that it is not the alternative of direct teaching learning.

It is a gentle request you to circulate this material only to your friends those who are studying in Semester IV (2020).

Stay safe and stay home.

Best wishes.

Paper GE 404:

Business Research Methods (BRM)

Module-I	
Unit 1: Introduction	Prof. Ram Prahlad Choudhary
Unit 3: Research Design	Mr. Pema Lama
Unit 4: Scaling Techniques and Questionnaire Design	Dr. Bappaditya Biswas
Unit 5: Sampling Design and Data Collection	Dr. Sajal Das
Module-II	
Unit 7: Univariate Data Analysis	Dr. Bikram Singh
Unit 8: Multivariate Data Analysis	Prof. Ashish Kumar Sana & Dr. Swapan Sarkar
Unit 9: Time Series Analysis	Professor (Dr.) S.S.Saha

GE 404: Business Research Methods

<u>Module – I</u>

<u>CHAPTER-1: INTRODUCTION</u>

In this Chapter, we shall study the basic concept of business research. Before describing the business research it would be wise to know about the Research and definition of research to get into insight of the business research. The whole chapter is organized as follows:

- Meaning and objective of Business Research,
- Types of Research,
- Importance of Research in Business Decision Making, and
- Ethics in Business Research

Research: Meaning

Dictionary definition of research is a careful investigation or inquiry specifically through search for new facts in any branch of knowledge. In simple words, research is an art of scientific investigation.

Research is defined as human activity based on intellectual application in the investigation of matter. The primary purpose for applied research is discovering, interpreting, and the development of methods and systems for the advancement of human knowledge.

"Research refers to the systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analysing the facts and reaching certain conclusions either in the form of solution(s) towards the concerned problem or in certain generalizations for some theoretical formulation." (Kothari and Garg)

According to Fred Kerlinger, research is an organized enquiry designed and carried out to provide information for solving problem.

Creswell says that "Research is a process of steps used to collect and analyze information to increase our understanding of a topic or issue".

Therefore, research is an original contribution to the existing stock of knowledge making for its advancement. It is pursuit of truth with the help of study, observation, comparison and experiment. Research is the search for knowledge through objective and systematic method of finding solution to a problem.

BUSINESS RESEARCH: MEANING AND OBJECTIVE

Meaning of Business Research:

Business research may be defined as research activities carried out relating to the different functionaries in the business and corporate world. Business research is a process of acquiring detailed information of all the areas of business and using such information in maximizing the sales and profit of the business. Such a study helps companies determine which product/service is most profitable or in demand. The definition of business research involves acquiring information and knowledge for professional or commercial purposes such as determining opportunities and goals for a business. An example of business research is gathering sales information and writing a detailed report on marketing and sales.

Business research is a systematic inquiry that provides information to guide managerial decisions. In other words, it is a process of planning, acquiring, analyzing, and disseminating relevant data, information, and insights to decision makers in ways that mobilize the organization to take appropriate actions that, in turn, maximize performance.

This entire process by which we attempt to solve problems is called research. Thus, research involves a series of well-thought-out and carefully executed activities that will enable the manager to know how organizational problems can be solved, or at least considerably minimized. Business research is an organized, systematic, data-based, critical, objective, scientific inquiry or investigation into a specific problem, undertaken with the purpose of finding answers or solutions to it. Fundamentally, research provides the needed information that guides managers to make informed decisions to successfully deal with problems.

When you run a business, there are several things you can research on. You research everything from market shares to sales. Business research helps you make intelligent and informed decisions and identify the key areas to invest your money in.

For example, an automobile company plans to unveil the latest car model in the market. For that, they need to develop strategies to explore and monitor customer demand. So, the company will conduct research to collect information and analyze market trends. This will help them draw better conclusions and come up with a fine quality car at the right price resulting in a larger market share.

In business, research is usually primarily conducted to resolve problematic issues in the areas of accounting, finance, management, and marketing. In **Accounting**, budget control systems, practices, and procedures are frequently examined. Inventory costing methods, accelerated depreciation, time-series behavior of quarterly earnings, transfer pricing, cash recovery rates, and taxation methods are some of the other areas that are researched. In **Finance**, the operations of financial institutions, optimum financial ratios, mergers and acquisitions, leveraged buyouts, inter corporate financing, yields on mortgages, the behavior of the stock exchange, etc become the focus of investigation. In **Management** research may cover the study of employee attitudes and behaviors, human resources management, the impact of changing demographics on management practices, production operations management, strategy formulation, information systems, etc. In **Marketing** research may address issues relating to product image, advertising, sales pro- motion, distribution, packaging, pricing, after-sales service, consumer preferences, new product development, and other marketing aspects.

Qualities of a Good Business Research:

Good business research generates dependable data that are derived by professionally conducted practices and that can be used reliably for business decision making. But, poor research is carelessly planned and conducted, resulting in data that a manager can't use to reduce his or her decision-making risks. Good business research follows the standards of the scientific method with the following qualities (characteristics):

1. ***Purpose clearly defined*:** The purpose of the business research—the problem involved or the decision to be made—should be clearly defined and sharply delineated in terms as unambiguous as possible. The statement of the decision problem should include its scope, its limitations, and the precise meanings of all words and terms significant to the research. Failure of the researcher to do this adequately may raise legitimate doubts in the minds of research report readers as to whether the researcher has sufficient understanding of the problem to make a sound proposal attacking it.
2. ***Research process detailed:*** The research procedures used should be described in sufficient detail to permit another researcher to repeat the research. This includes the steps to acquire participants, informed consent, sampling methods and representativeness, and data gathering procedures. Omission of significant procedural details makes it difficult or impossible to estimate the validity and reliability of the data and justifiably weakens the confidence of the reader in the research itself as well as any recommendations based on the research.
3. ***Research design thoroughly planned:*** The procedural design of the research, and its choice among competing designs, should be clearly described and carefully planned to yield results that are as objective as possible. A survey of opinions or recollections ought not to be used when more reliable evidence is available from documentary sources or by direct observation. Bibliographic searches should be as thorough and complete as possible. Experiments should have satisfactory controls, reducing threats to internal validity and enhancing the probability of external validity (generalizability). Direct

observations should be recorded as soon as possible after the event. Efforts should be made to minimize the influence of personal bias in selecting and recording data.

4. ***High ethical standards applied***: Researchers often work independently and have significant latitude in designing and executing projects. A research design that includes safeguards against causing mental or physical harm to participants and makes data integrity a first priority should be highly valued. Ethical issues in research reflect important moral concerns about the practice of responsible behavior in society. Careful consideration must be given to those research situations in which there is a possibility for physical or psychological harm, exploitation, invasion of privacy, and/or loss of dignity. The research need must be weighed against the potential for these adverse effects. Typically, you can redesign a study, but sometimes you cannot. The researcher should be prepared for this dilemma.
5. ***Limitations frankly revealed:*** The researcher should report, with complete frankness, flaws in procedural design and estimate their effect on the findings. There are very few perfect research designs. Some of the imperfections may have little effect on the validity and reliability of the data; others may invalidate them entirely. A competent researcher should be sensitive to the effects of imperfect design. The researcher's experience in analyzing data should provide a basis for estimating the influence of design flaws. As a decision maker, you should question the value of research about which no limitations are reported.
6. ***Adequate analysis for decision maker's needs***: Analysis of the data should be extensive enough to reveal its significance, what managers call insights. The methods of analysis used should be appropriate. The extent to which this criterion is met is frequently a good measure of the competence of the researcher. Adequate analysis of the data is the most difficult phase of research for the novice. The validity and reliability of data should be checked carefully. The data should be classified in ways that assist the researcher in reaching pertinent conclusions and clearly reveal the findings that have led to those conclusions. When statistical methods are used, appropriate descriptive and inferential techniques should be chosen, the probability of error should be estimated, and the criteria of statistical significance applied.
7. ***Findings presented unambiguously***: Some evidence of the competence and integrity of the researcher may be found in the report itself. For example, language that is restrained, clear, and precise; assertions that are carefully drawn and hedged with appropriate reservations; and an apparent effort to achieve maximum objectivity tend to leave a favorable impression of the researcher with the decision maker. Presentation of data should be comprehensive, reasonably interpreted, easily understood by the decision maker, and organized so that the decision maker can readily locate critical findings.

8. ***Conclusions justified***: Conclusions should be limited to those for which the data provide an adequate basis. Researchers are often tempted to broaden the basis of induction by including personal experiences and their interpretations—data not subject to the controls under which the research was conducted. Equally undesirable is the all-too-frequent practice of drawing conclusions from a study of a limited population and applying them universally. Good researchers always specify the conditions under which their conclusions seem to be valid.
9. ***Researcher's experience reflected***: Greater confidence in the research is warranted if the researcher is experienced, has a good reputation in research, and is a person of integrity. Were it possible for the reader of a research report to obtain sufficient information about the researcher, this criterion perhaps would be one of the best bases for judging the degree of confidence a piece of research warrants and the value of any decision based upon it. For this reason the research report should contain information about the qualifications of the researcher.

Good business research has an inherent value only to the extent that it helps management make better decisions to achieve organizational goals. Interesting information about consumers, employees, competitors, or the environment might be pleasant to have, but its value is limited if the information cannot be applied to a critical decision. If a study does not help management select more effective, more efficient, less risky, or more profitable alternatives than otherwise would be the case, its use should be questioned. Alternatively, management may have insufficient resources (time, money, or skill) to conduct an appropriate study or may face a low level of risk associated with the decision at hand. In these situations, it is valid to avoid business research and its associated costs in time and money. Business research finds its justification in the contribution it makes to the decision maker's task and to the bottom line.

Objective of Business Research:

The ultimate aim of research is to find (search) answer to questions by applying a scientific research process. Although, there are different research objectives for different research studies, some general objectives may be mentioned hereunder:

1. To gain familiarity with a phenomenon or to achieve new insights into it (exploratory or formulative research studies)
2. To describe accurately the characteristics of a particular individual, situation or a group. (descriptive research)
3. To determine the frequency with which something occurs or with which it is associated with something else. (studies with this object known as diagnostic research)

4. To test a hypothesis of a causal relationship between variables. (such studies are known as hypothesis testing research)

TYPES OF RESEARCH

The essential types of research are as follows:

1. Basic Research:

- It is also known as pure or fundamental research.
- This research is mainly conducted to increase knowledge base. It is driven purely by interest and a desire to expand our knowledge.
- This type of research tends not to be directly applicable to the real world in a direct way, but enhances our understanding of the world around us.
- Pure research can be exploratory, descriptive or explanatory.
- Basic research generates new ideas, principles and theories indifferent fields.
- Basic research concentrates on fundamental principles and testing theories.
- It is sometimes implicitly said that basic research doesn't have practical applications. For example, someone conducting basic research on cheating behavior may design a study examining whether students from illiterate families cheat more often than students from literate families.
- Notice that the research is not done to reduce cheating or help people who cheat or any other "applied" aspect, but to increase the understanding of cheating behavior.

2. Applied Research:

- Applied research is mainly related with solving practical problems rather than focusing on knowledge expansion.
- It is mainly used to find solutions to problems which occur on a daily basis and develop new innovative technologies.
- The main aim of applied research is to provide better technologies for humans to enhance their standard of living.
- Example: Investigating which treatment approach is the most effective for treating cancer patients whereas researching which strategies work best to motivate workers.

3. Quantitative Research:

- Quantitative research is generally related with the positivist concept.
- It usually involves collecting and converting data into numerical form so that statistical calculations can be made and conclusions drawn.
- Objectivity is very vital in quantitative research.

- Therefore, researchers try to avoid their own presence, behavior or attitude affecting the results (e.g., by changing the circumstances being studied or causing participants to behave differently).
- They also examine their methods and results for any possible bias.
- The aim of quantitative research is to develop mathematical models, theories related to phenomenon. Quantitative research is mainly used in social sciences.

4. Qualitative Research:

- Qualitative research is the approach usually related with the social constructivist concept which emphasizes the socially constructed nature of reality.
- It is about recording, analyzing and attempting to reveal the in debt meaning and significance of human behavior and experience, including conflicting beliefs, behaviors and emotions.
- The qualitative method tries to answer why and how of decision-making rather than what and when.
- The approach to data collection and analysis is logical but allows for greater flexibility than in quantitative research.
- Data is collected in textual form on the basis of observation and communication with the participants, e.g. through participant observation, in-depth interviews and focus groups.
- It is not converted into numerical form and is not statistically analyzed.

5. Descriptive Research:

- Descriptive research is used to describe characteristics of an observable fact being studied.
- Descriptive studies are structured in such a way that it cannot be changed frequently, so it can be said that they are rigid in nature.
- They cannot identify cause and effect relationship between variables.
- Descriptive research answers questions such as who, when, where, what and how.
- This type of research describes what exists and may help to reveal new facts and meaning.
- The purpose of descriptive research is to observe, describe and document.

6. Exploratory Research:

- Exploratory research is carried out for a problem that has not been clearly defined.
- The main aim of this research is to gather initial information which helps to define problems and recommend hypothesis.

- Exploratory research helps to settle on the best research design, data collection method and selection of subjects.
- Exploratory research often relies on secondary research such as reviewing available literature, or qualitative approaches such as informal discussions with consumers, employees, management or competitors, and more formal approaches through in-depth interviews, focus groups, projective methods, case studies or pilot studies.
- Exploratory research can mainly be conducted when researchers lack clear idea of the problem.
- The results of exploratory research are not generally useful for decision-making, but they can provide major insight into a given situation.

7. Historical Research:

- It is defined as the type of research that examines past events or combinations of events to arrive at an account of what has happened in the past.
- Historical research is carried out to discover the unknown; answer questions, recognize the relationship that the past has to the present; record and assess activities of individuals, agencies, or institutions; and assist in understanding the culture in which we live.
- Historical research can exhibit patterns that occurred in the past and over time which can facilitate us to see where we came from and what kinds of solutions we have used in the past.
- We usually will notice that what we do today is expressly rooted in the past. Historical research involves the process of collecting and reading the research material collected, and writing the document from the data collected.

8. Experimental Research:

- It is commonly used in sciences such as sociology and psychology, physics, chemistry, biology, medicine, etc.
- It is a collection of research designs which use manipulation and controlled testing to understand fundamental processes.
- Usually, one or more variables are manipulated to establish their effect on a dependent variable.
- Experimental Research is mainly used when: there is time priority in a causal relationship (cause precedes effect) or there is uniformity in a causal relationship (a cause will always lead to the same effect) or the magnitude of the correlation is great.
- Experimental research is important to society as it helps us to improve our daily lives.

(***Note:*** Researchers use various research methods to collect relevant data so that business enterprises can make wiser decisions. There are two main types of methods to carry out business research.

- **Quantitative Business Research**
  It is a method of analyzing the largest group that meets your target goals. It uses mathematical techniques and data to explain the important stats about your business and market. Usually, this data uses multiple-choice questionnaires that can help you be profitable with your sales. For instance, quantitative research can answer questions such as;
- Are your customers aware of the services or products you offer?
- How many people are interested in buying your products or services?
- Who are your best customers and what are their buying habits?
- How long the visitor stays on your website, and which is their exit page?

The result of quantitative business research is in the numerical form, such as;

- 40% of customers rate the new product as "attractive"
- 70% of prospective customers use the Internet to book their hotel room
- 6 out of 10 customers will buy a new food product after trying the free in-store sample

The quantitative research methods include various surveys such as postal, telephone, online, and face-to-face.

- **Qualitative Business Research**

This business research focuses on attitudes, intentions, and beliefs. Qualitative research includes questions such as "Why"? or "How?".

The aim of this research is to gain insights into customers' distinct behaviors and response to a new product. This research is beneficial for your new products and marketing initiatives to test reactions and rectify your approach.

You can collect qualitative data using common methods such as case studies, focus groups, and interviews. This data is often valuable but can be time-consuming and expensive to collect, especially for a small business or a startup.)

IMPORTANCE OF RESEARCH IN BUSINESS DECISION MAKING

Business research helps businesses understand their customers' buying patterns, preferences and pain points, gain deeper insights into the contenders, current market trends, and demographics. Using effective strategies to understand the demand and supply of the market, businesses can always stay ahead of the competition. Using business research, they can reduce costs and design solutions that aim at the market demand and their target audience.

In business chances of failures are less with business research as it gives an idea of the target customers and the perfect time to launch a product. In addition, with a deep understanding of brand value, businesses can constantly innovate to meet customer requirements. This is essential to grow market share and revenue. The SWOT analysis in business research is crucial to make an informed decision and making the business a huge success.

Research is the building block of any business. It acts as a catalyst to thrive in the market. So, never underestimate the value of market research and leverage its benefits to give an extra edge to your business.

Besides the above the following are the importance of Business Research:

1. A research problem refers to a complexity which a researcher or a scientific community or an industry or a government organization or a society experiences. It may be a theoretical or a practical situation. It calls for a systematic understanding and possible solution.

2. Research on existing theories and concepts help us recognize their range and applications.

3. It is the bank of knowledge and provides strategy for solving problems.

4. It is important in industry and business for higher profits, output, efficiency and to improve the quality of products.

5. Mathematical and logical research on business and industry reduces the problems in them.

6. It leads to the identification and categorization of new materials, new living things, new stars, etc.

7. Inventions can be done through research 8. Social research helps find answers to social problems. They explain social phenomena and try to find solution to social problems.

ETHICS IN BUSINESS RESEARCH

Ethics in business research refers to a code of conduct or expected societal norm of behavior while conducting research. Ethical conduct applies to the organization and the members that sponsor the research, the researchers who undertake the research, and the respondents who provide them with the necessary data. The observance of ethics begins with the person instituting the research, who should do so in good faith, pay attention to what the results indicate, and surrendering the ego, pursue organizational rather than self-interests. Ethical conduct should also be reflected in the behavior of the researchers who conduct the investigation, the participants who provide the data, the analysts who provide the results, and the entire research team that presents the interpretation of the results and suggests alternative solutions.

Thus, ethical behavior pervades each step of the research process—data collection, data analysis, reporting, and dissemination of information on the Internet, if such an activity is undertaken. How the subjects are treated and how confidential information is safeguarded are all guided by business ethics. The American Psychological Association has established certain guidelines for conducting research, to ensure that organizational research is conducted in an ethical manner and the interests of all concerned are safeguarded.

As in other aspects of business, all parties in research should exhibit ethical behavior. Ethics are norms or standards of behavior that guide moral choices about our behavior and our relationships with others. The goal of ethics in research is to ensure that no one is harmed or suffers adverse consequences from research activities. This objective is usually achieved. However, unethical activities are pervasive and include violating nondisclosure agreements, breaking participant confidentiality, misrepresenting results, deceiving people, using invoicing irregularities, avoiding legal liability, and more.

The recognition of ethics as a problem for economic organizations is repeatedly revealed in surveys. Despite an increase in awareness of formal ethics programs and the presence of written ethical codes of conduct there is a reporting of ethical misconduct. There is no single approach to ethics. Advocating strict adherence to a set of laws is difficult because of the unforeseen constraint put on researchers. Alternatively, relying on each individual's personal sense of morality is equally problematic.

Consider the clash between those who believe death is deliverance from a life of suffering and those who value life to the point of preserving it in definitely through mechanical means. Each value system claims superior knowledge of moral correctness. Clearly, a middle ground between being completely code governed or relying on ethical relativism is necessary. The foundation for that middle ground is an emerging consensus on ethical standards for researchers. Codes and regulations guide researchers and sponsors. Review boards and peer groups help researchers examine their research proposals for ethical dilemmas.

Many design-based ethical problems can be eliminated by careful planning and constant vigilance. In the end, responsible research anticipates ethical dilemmas and attempts to adjust the design, procedures, and protocols during the planning process rather than treating them as an afterthought. Ethical research requires personal integrity from the researcher, the project manager, and the research sponsor. Integrity in research is vital.

***

<u>Multiple Choice Questions:</u>

1. Research is pursuit of truth with the help of study, observation, comparison and
    a. regulation
    b. experiment
    c. entertainment
    d. differentiation

2. Good business research generates dependable data that are used reliably for
    a. **business decision making**
    b. national policy making
    c. international policy making
    d. profit making
3. Which of the following is ***not*** a quality of a good business research?
    a. Purpose clearly defined
    b. High ethical standards applied
    c. **Highly complex and sophisticated**
    d. Research process detailed
4. Basic research is otherwise known as
    a. empirical research
    b. **fundamental research**
    c. descriptive research
    d. qualitative research
5. Ethical behavior pervades each step of the
    a. report writing
    b. data collection
    c. **research process**
    d. data analysis

References:

- Kothari, C R and Garg, Gaurav, Research Methodology- Methods and Techniques, New Age International Publishers, 2014
- Mishra, Prahlad, Business Research Methods, Oxford University Press, New Delhi, 2015
- Cooper, Donald R. and Schindler, Pamela S, Business Research Methods, McGraw-Hill/Irwin, New York, 2014
- Bhome, S. S, Prajapati, N, Banga, C. K, and Desai, Swati S, Business Research Methods, Himalaya Publishing House Pvt. Ltd., Mumbai, 2016

****

Lecture Note Prepared By:

Prof. Ram Prahlad Choudhary

Dept. of Commerce, University of Calcutta

Module – I:
Unit -3 (Research Design)

A research design is a systematic plan to study a scientific problem. The design of a study defines the study type (descriptive, correlation, semi-experimental, experimental, review, meta-analytic) and sub-type (e.g., descriptive - longitudinal case study), research question, hypotheses, independent and dependent variables, experimental design, and, if applicable, data collection methods and a statistical analysis plan. Research design is the framework that has been created to seek answers to research questions.

Need and Importance of Research Design

Research design carries an important influence on the reliability of the results attained. It therefore provides a solid base for the whole research. It is needed due to the fact that it allows for the smooth working of the many research operations. This makes the research as effective as possible by providing maximum information with minimum spending of effort, money and time. For building of a car, we must have a suitable blueprint made by an expert designer. In a similar fashion, we require a suitable design or plan just before data collection and analysis of the research project. Planning of design must be carried out cautiously as even a small mistake might mess up the purpose of the entire project. The design helps the investigator to organize his ideas, which helps to recognize and fix his faults, if any.

Experimental Research Design

A blueprint of the procedure that enables the researcher to maintain control over all factors that may affect the result of an experiment. In doing this, the researcher attempts to determine or predict what may occur. Experimental research is often used where there is time priority in a causal relationship (cause precedes effect), there is consistency in a causal relationship (a cause will always lead to the same effect) and the magnitude of the correlation is great. The classic experimental design specifies an experimental group and a control group. The independent variable is administered to the experimental group and not to the control group, and both groups are measured on the same dependent variable. Subsequent experimental designs have used more groups and more measurements over longer periods. True experiments must have control, randomization and manipulation.

Cross - Sectional Research Design

The cross-sectional design can only measure differences between or from among a variety of people, subjects, or phenomena rather than a process of change. As such, researchers using this design can only employ a relatively passive approach to making causal inferences based on findings.

Quantitative versus Qualitative Research Design

Quantitative Research Design	Qualitative Research Design
• The aim is to classify features, count them, and construct statistical models in an attempt to explain what is observed	• The aim is a complete, detailed description
• The researcher knows clearly in advance what he/she is looking for	• The researcher may only know roughly in advance what he/she is looking for
• Recommended during latter phases of research projects	• Recommended during earlier phases of research projects
• All aspects of the study are carefully designed before data is collected	• The design emerges as the study unfolds
• The researcher uses tools such as questionnaires or equipment to collect numerical data	• The researcher is the data gathering instrument
• Data are in the form of numbers and statistics	• Data are in the form of words, pictures or objects
• Objective – seeks precise measurement and analysis of target concepts, e.g. uses surveys, questionnaires etc.	• Subjective – individuals' interpretation of events is important, e.g. uses participant observation, in-depth interviews etc.
• Quantitative data are more efficient, able to test hypotheses, but may miss contextual detail	• Qualitative data are more rich, time consuming, and less able to be generalized
(Adapted from Neill, 2007)	

Descriptive Research Design

It help to provide answers to the questions of who, what, when, where, and how associated with a particular research problem; a descriptive study cannot conclusively ascertain answers to why. Descriptive research is used to obtain information concerning the current status of the phenomena and to describe "what exists" with respect to variables or conditions in a situation.

Research design in case of descriptive and diagnostic research studies

Descriptive research studies are those studies which are concerned with describing the characteristics of a particular individual, or of a group, whereas diagnostic research studies determine the frequency with which something occur or its association with something else. The studies concerning whether certain variables are associated are examples of diagnostic research studies. As against this, studies concerned with specific predictions, with narration of facts and characteristics concerning individual, group or situation are all examples of descriptive research studies. Most of the social research comes under this category. In descriptive as well as in diagnostic studies, the researcher must be able to define clearly, what he wants to measure and must find adequate methods for measuring it along with a clear cut definition of 'population' he wants to study. Since the aim is to obtain complete and accurate information in the said studies, the procedure to be used must be carefully planned. The research design must make enough provision for protection against bias and must maximize reliability, with due concern for the economical completion of the research study. The design in such studies must be rigid and not flexible and must focus attention on the following:

- Formulating the objective of the study *(what the study is about and why is it being made?)*
- Designing the methods of data collection *(what techniques of gathering data will be adopted?)*
- Selecting the sample *(how much material will be needed?)*
- Collecting the data *(where can the required data be found and with what time period should the data be related?)*
- Processing and analysing the data.
- Reporting the findings.

Longitudinal Research Design

It follows the same sample over time and makes repeated observations. For example, with longitudinal surveys, the same group of people is interviewed at regular intervals, enabling researchers to track changes over time and to relate them to variables that might explain why the changes occur. Longitudinal research designs describe patterns of change and help establish the direction and magnitude of causal relationships. Measurements are taken on each variable over two or more distinct time periods. This allows the researcher to measure change in variables over time. It is a type of observational study sometimes referred to as a panel study.

Causal Research Design

Causality studies may be thought of as understanding a phenomenon in terms of conditional statements in the form, "If X, then Y." This type of research is used to measure what impact a specific change will have on existing norms and assumptions. Most social scientists seek causal explanations that reflect tests of hypotheses. Causal effect (nomothetic perspective) occurs when variation in one phenomenon, an independent variable, leads to or results, on average, in variation in another phenomenon, the dependent variable.

Conditions necessary for determining causality

• **Empirical association:** a valid conclusion is based on finding an association between the independent variable and the dependent variable.

• **Appropriate time order:** to conclude that causation was involved, one must see that cases were exposed to variation in the independent variable before variation in the dependent variable.

• **Non-spuriousness:** a relationship between two variables that is not due to variation in a third variable.

Exploratory Research Design

It is conducted about a research problem when there are few or no earlier studies to refer to or rely upon to predict an outcome. The focus is on gaining insights and familiarity for later investigation or undertaken when research problems are in a preliminary stage of investigation. These designs are often used to establish an understanding of how best to proceed in studying an issue or what methodology would effectively apply to gathering information about the issue.

Research design in case of Exploratory Research Studies

The major emphasis in such studies is on the discovery of ideas and insights. Generally, the following three methods in the context of research design for such studies are as follows:

(a) The survey of concerning literature: It happens to be the most simple and fruitful method of formulating precisely the research problem or developing hypothesis. Hypotheses stated by earlier workers may be reviewed and their usefulness be evaluated as a basis for further research. It may also be considered whether the already stated hypotheses suggest new hypothesis. In this way the researcher should review and build upon the work already done by others, but in cases where hypotheses have not yet been formulated, his task is to review the available material for deriving the relevant hypotheses from it. Besides, the bibliographical survey of studies, already made in one's area of interest may as well as made by the researcher for precisely formulating the problem. He should also make an attempt to apply concepts and theories developed in different research contexts to the area in which he is himself working. Sometimes the works of creative writers also provide a fertile ground for hypothesis formulation and as such may be looked into by the researcher.

(b) The Experience survey: It means the survey of people who have had practical experience with the problem to be studied. The object of such a survey is to obtain insight into the relationships between variables and new ideas relating to the research problem. For such a survey people who are competent and can contribute new ideas may be carefully selected as respondents to ensure a representation of different types of experience. The respondents so selected may then be interviewed by the investigator. The researcher must prepare an interview schedule for the systematic questioning of informants. But the interview must ensure flexibility in the sense that the respondents should be allowed to raise issues and questions which the investigator has not previously considered. Generally, the experience collecting interview is likely to be long and may last for few hours. Hence, it is often considered desirable to send a copy of the questions to be discussed to the respondents well in advance. This will also give an opportunity to the respondents for doing some advance thinking over the various issues involved so that, at the time of interview, they may be able to contribute effectively. Thus, an experience survey may enable the researcher to define the problem more concisely and help in the formulation of the research hypothesis. This survey may as well provide information about the practical possibilities for doing different types of research.

(c) The analysis of 'insight-stimulating' examples: It is also a fruitful method for suggesting hypotheses for research. It is particularly suitable in areas where there is little experience to serve as a guide. This method consists of the intensive study of selected instances of the phenomenon in which one is interested. For this purpose the existing records, if any, may be examined, the unstructured interviewing may take place, or some other approach may be adopted. Attitude of the investigator, the intensity of the study and the ability of the researcher to draw together diverse information into a unified interpretation are the main features which make this method an appropriate procedure for evoking insights.

Thus, in an exploratory of formulative research study which merely leads to insights or hypotheses, whatever method or research design outlined above is adopted, the only thing essential is that it must continue to remain flexible so that many different facets of a problem may be considered as and when they arise and come to the notice of the researcher.

Suggested Readings

- **C. R. Kothari and Gaurav Garg**: Research Methodology (Methods and Techniques) New Age International Publishers
- **Prahlad Mishra**: Business Research Methods Oxford University Press

Paper GE.404: Business Research Methods

Dr. Bappaditya Biswas

Module I

Topic: Scaling Techniques and Questionnaire Design

1. Concept of Measurement

Measurement means assigning number or other symbols to characteristics of objects according to certain pre-specified rules. In our daily life we are said to measure when we use some yardstick to determine weight, height, or some other feature of a physical object. We also measure when we judge how well we like a song, a painting or the personalities of our friends. We, thus, measure physical objects as well as abstract concepts. Measurement is a relatively complex and demanding task, specially so when it concerns qualitative or abstract phenomena. By measurement we mean the process of assigning numbers to objects or observations, the level of measurement being a function of the rules under which the numbers are assigned.

It is easy to assign numbers in respect of properties of some objects, but it is relatively difficult in respect of others. For instance, measuring such things as social conformity, intelligence, or marital adjustment is much less obvious and requires much closer attention than measuring physical weight, biological age or a person's financial assets. In other words, properties like weight, height, etc., can be measured directly with some standard unit of measurement, but it is not that easy to measure properties like motivation to succeed, ability to stand stress and the like. We can expect high accuracy in measuring the length of pipe with a yard stick, but if the concept is abstract and the measurement tools are not standardized, we are less confident about the accuracy of the results of measurement.

2. Measurement Scales

From what has been stated above, we can write that scales of measurement can be considered in terms of their mathematical properties. The most widely used classification of measurement scales are: (a) nominal scale; (b) ordinal scale; (c) interval scale; and (d) ratio scale.

(a) Nominal scale:

Nominal scale is simply a system of assigning number symbols to events in order to label them. The usual example of this is the assignment of numbers of basketball players in order to identify them. Such numbers cannot be considered to be associated with an ordered scale for their order is of no consequence; the numbers are just convenient labels for the particular class of events and as such have no quantitative value. Nominal scales provide convenient ways of keeping track of people, objects and events. One cannot do much with the numbers involved.

For example, one cannot usefully average the numbers on the back of a group of football players and come up with a meaningful value. Neither can one usefully compare the numbers assigned to one group with the numbers assigned to another. The counting of members in each group is the only possible arithmetic operation when a nominal scale is employed. Accordingly, we are restricted to use mode as the measure of central tendency. There is no generally used measure of dispersion for nominal scales. Chi-square test is the most common test of statistical significance that can be utilized, and for the measures of correlation, the contingency coefficient can be worked out.

Nominal scale is the least powerful level of measurement. It indicates no order or distance relationship and has no arithmetic origin. A nominal scale simply describes differences between things by assigning them to categories. Nominal data are, thus, counted data. The scale wastes any information that we may have about varying degrees of attitude, skills, understandings, etc. In spite of all this, nominal scales are still very useful and are widely used in surveys and other *ex-post-facto* research when data are being classified by major sub-groups of the population.

Characteristics

- It has no arithmetic origin.
- It shows no order or distance relationship.
- It distinguishes things by putting them into various groups.

Use

This scale is generally used in conducting in surveys and ex-post-facto research.

Example:

Have you ever visited Bangalore?

Yes-1

No-2

'Yes' is coded as 'One' and 'No' is coded as 'Two'. The numeric attached to the answers has no meaning, and is a mere identification. If numbers are interchanged as one for 'No' and two for 'Yes', it won't affect the answers given by respondents. The numbers used in nominal scales serve only the purpose of counting.

The telephone numbers are an example of nominal scale, where one number is assigned to one subscriber. The idea of using nominal scale is to make sure that no two persons or objects receive the same number. Similarly, bus route numbers are the example of nominal scale. "How old are you"? This is an example of a nominal scale.

Limitations

- There is no rank ordering.
- No mathematical operation is possible.
- Statistical implication - Calculation of the standard deviation and the mean is not possible. It is possible to express the mode.

(b) Ordinal scale:

The lowest level of the ordered scale that is commonly used is the ordinal scale. The ordinal scale places events in order, but there is no attempt to make the intervals of the scale equal in terms of some rule. Rank orders represent ordinal scales and are frequently used in research relating to qualitative phenomena. A student's rank in his graduation class involves the use of an ordinal scale. One has to be very careful in making statement about scores based on ordinal scales.

For instance, if Ram's position in his class is 10 and Mohan's position is 40, it cannot be said that Ram's position is four times as good as that of Mohan. The statement would make no sense at all. Ordinal scales only permit the ranking of items from highest to lowest. Ordinal measures have no absolute values, and the real differences between adjacent ranks may not be equal. All that can be said is that one person is higher or lower on the scale than another, but more precise comparisons cannot be made.

Thus, the use of an ordinal scale implies a statement of 'greater than' or 'less than' (an equality statement is also acceptable) without our being able to state how much greater or less. The real difference between ranks 1 and 2 may be more or less than the difference between ranks 5 and 6. Since the numbers of this scale have only a rank meaning, the appropriate measure of central tendency is the median. A percentile or quartile measure is used for measuring dispersion. Correlations are restricted to various rank order methods. Measures of statistical significance are restricted to the non-parametric methods.

Characteristics

- The ordinal scale ranks the things from the highest to the lowest.
- Such scales are not expressed in absolute terms.
- The difference between adjacent ranks is not equal always.
- For measuring central tendency, median is used.
- For measuring dispersion, percentile or quartile is used.

Example:

The respondents may be given a list of brands which may be suitable and were asked to rank on the basis of ordinal scale:

- Lux
- Liril
- Cinthol
- Lifebuoy
- Park Avenue

(c) Interval scale:

In the case of interval scale, the intervals are adjusted in terms of some rule that has been established as a basis for making the units equal. The units are equal only in so far as one accepts the assumptions on which the rule is based. Interval scales can have an arbitrary zero, but it is not possible to determine for them what may be called an absolute zero or the unique origin. The primary limitation of the interval scale is the lack of a true zero; it does not have the capacity to measure the complete absence of a trait or characteristic. The Fahrenheit scale is an example of an interval scale and shows similarities in what one can and cannot do with it. One can say that an increase in temperature from 30° to 40° involves the same increase in temperature as an increase from 60° to 70°, but one cannot say that the temperature of 60° is twice as warm as the temperature of 30° because both numbers are dependent on the fact that the zero on the scale is set arbitrarily at the temperature of the freezing point of water. The ratio of the two temperatures, 30° and 60°, means nothing because zero is an arbitrary point.

Interval scales provide more powerful measurement than ordinal scales for interval scale also incorporates the concept of equality of interval. As such more powerful statistical measures can be used with interval scales. Mean is the appropriate measure of central tendency, while standard deviation is the most widely used measure of dispersion. Product moment correlation techniques are appropriate and the generally used tests for statistical significance are the ‘t’ test and ‘F’ test.

Characteristics

- Interval scales have no absolute zero. It is set arbitrarily.
- For measuring central tendency, mean is used.
- For measuring dispersion, standard deviation is used.
- For test of significance, t-test and f-test are used.
- Scale is based on the equality of intervals.

Use

Most of the common statistical methods of analysis require only interval scales in order that they might be used. These are not recounted here because they are so common and can be found in virtually all basic texts on statistics.

Interval scales may be either numeric or semantic.

Suppose we want to measure the rating of a refrigerator using interval scale by scoring them on a scale of 5 down to 1 (i.e. 5 = Excellent; 1= Poor) on each of the criteria listed. Circle the appropriate score on each line.

Brand name	5	4	3	2	1
Price	5	4	3	2	1
After Sale Service	5	4	3	2	1
Utility	5	4	3	2	1
Attractively Design	5	4	3	2	1

The researcher cannot conclude that the respondent who gives a rating of 4 is 2 times more favourable towards a product under study than another respondent who awards the rating of 2

Please indicate your views on Jack Olive Oil by ticking the appropriate responses below:

	Excellent	Very Good	Good	Fair	Poor
Succulent					
Freshness					
Freedom from skin blemish					
Value for money					
Attractiveness of packaging					

d. Ratio scale:

The highest level of measurement is a ratio scale. This has the properties of an interval scale together with a fixed origin or zero point. Examples of variables which are ratio scaled include weights, lengths and times. Ratio scales permit the researcher to compare both differences in scores and the relative magnitude of scores. For instance the difference between 5 and 10 minutes is the same as that between 10 and 15 minutes, and 10 minutes is twice as long as 5 minutes.

Ratio scales have an absolute or true zero of measurement. The term 'absolute zero' is not as precise as it was once believed to be. We can conceive of an absolute zero of length and similarly we can conceive of an absolute zero of time. For example, the zero point on a centimeter scale indicates the complete absence of length or height. But an absolute zero of temperature is theoretically unobtainable and it remains a concept existing only in the scientist's mind. The number of minor traffic-rule violations and the number of incorrect letters in a page of type script represent scores on ratio scales. Both these scales have absolute zeros and as such all minor traffic violations and all typing errors can be assumed to be equal in significance. With ratio scales involved one can make statements like "Jyoti's" typing performance was twice as good as that of "Reetu." The ratio involved does have significance and facilitates a kind of comparison which is not possible in case of an interval scale.

Ratio scale represents the actual amounts of variables. Measures of physical dimensions such as weight, height, distance, etc. are examples. Generally, all statistical techniques are usable with ratio scales and all manipulations that one can carry out with real numbers can also be carried out with ratio scale values. Multiplication and division can be used with this scale but not with other scales mentioned above. Geometric and harmonic means can be used as measures of central tendency and coefficients of variation may also be calculated.

Thus, proceeding from the nominal scale (the least precise type of scale) to ratio scale (the most precise), relevant information is obtained increasingly. If the nature of the variables permits, the researcher should use the scale that provides the most precise description. Researchers in physical sciences have the advantage to describe variables in ratio scale form but the behavioural sciences are generally limited to describe variables in interval scale form, a less precise type of measurement.

Characteristics

- This scale has an absolute zero measurement.
- For measuring central tendency, geometric and harmonic means are used.

Use:

Ratio scale can be used in all statistical techniques.

Example: Sales this year for product A are twice the sales of the same product last year.

Summary

Type	**Basic empirical operation**	**Typical usage**	**Typical Statistics Descriptive**	**Inferential**
1. Nominal	Determination Of equality (0, 1,2, 9)	Classification Male-female purchaser non-purchaser, Team A Team-B	Percentage, Mode	Chi-square, Binomial test
2. Ordinal	Determination Of greater or less (o<1<2… <9)	Rankings: preference data, market position, attitude measures, many psychological measures	Percentile, Median	Rank-order correlation
3. Interval	Determination of equality of intervals (2-1=7-6)	Index numbers, attitude measures	Mean, Range, Standard Deviation	Product-moment correlation
4. Ratio	Determination of equality of ratios (2/4 = 4/8)	Sales, units produced, number	Geometric mean	Coefficient of variation

		of customers. Costs, age		

3. Scaling

In research we quite often face measurement problem (since we want a valid measurement but may not obtain it), specially when the concepts to be measured are complex and abstract and we do not possess the standardised measurement tools. Alternatively, we can say that while measuring attitudes and opinions, we face the problem of their valid measurement.

Scaling is a process or set of procedures, which is used to assess the attitude of an individual. Scaling is defined as the assignment of objects to numbers according to a rule. The objects in the definition are text statements, which can be the statements of attitude or principle. Attitude of an individual is not measured directly by scaling. It is first migrated to statements and then the numbers are assigned to them.

Meaning of Scaling

Scaling describes the procedures of assigning numbers to various degrees of opinion, attitude and other concepts. This can be done in two ways viz., (i) making a judgement about some characteristic of an individual and then placing him directly on a scale that has been defined in terms of that characteristic and (ii) constructing questionnaires in such a way that the score of individual's responses assigns him a place on a scale.

It may be stated here that a scale is a continuum, consisting of the highest point (in terms of some characteristic e.g., preference, favourableness, etc.) and the lowest point along with several intermediate points between these two extreme points. These scale-point positions are so related to each other that when the first point happens to be the highest point, the second point indicates a higher degree in terms of a given characteristic as compared to the third point and the third point indicates a higher degree as compared to the fourth and so on. Numbers for measuring the distinctions of degree in the attitudes/opinions are, thus, assigned to individuals corresponding to their scale-positions. All this is better understood when we talk about scaling technique(s). Hence the term 'scaling' is applied to the procedures for attempting to determine quantitative measures of subjective abstract concepts.

Scaling has been defined as a "procedure for the assignment of numbers (or other symbols) to a property of objects in order to impart some of the characteristics of numbers to the properties in *question.”*

4. Different Scaling Techniques

We now take up some of the important scaling techniques often used in the context of research specially in context of social or business research.

(i) Rating scales:

The rating scale involves qualitative description of a limited number of aspects of a thing or of traits of a person. When we use rating scales (or categorical scales), we judge an object in absolute terms against some specified criteria i.e., we judge properties of objects without reference to other similar objects. These ratings may be in such forms as “like-dislike”, “above average, average, below average”, or other classifications with more categories such as “like very much—like some what—neutral—dislike somewhat—dislike very much”; “excellent—good—average—below average—poor”, “always—often—occasionally—rarely—never”, and so on. There is no specific rule whether to use a two-points scale, three-points scale or scale with still more points. In practice, three to seven points scales are generally used for the simple reason that more points on a scale provide an opportunity for greater sensitivity of measurement.

The following rating scales are often used in organizational research:

a. Dichotomous scale
b. Category scale
c. Likert scale
d. Numerical scales
e. Semantic differential scale
f. Itemized rating scale
g. Fixed or constant sum rating scale
h. Stapel scale
i. Graphic rating scale
j. Consensus scale

Other scales such as the Thurstone Equal Appearing Interval Scale, and the Multidimensional Scale are less frequently used. We will briefly describe each of the above attitudinal scales.

a. Dichotomous Scale

The dichotomous scale is used to elicit a Yes or No answer, as in the example below. Note that a nominal scale is used to elicit the response.

Example

Do you own a car? Yes No

b. Category Scale

The category scale uses multiple items to elicit a single response as per the fol- lowing example. This also uses the nominal scale.

Example

Where in northern California do you reside? North Bay
South Bay East
Bay Peninsula
Other

c. Likert Scale

The Likert scale is designed to examine how strongly subjects agree or disagree with statements on a 5-point scale with the following anchors:

Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
1	2	3	4	5

The responses over a number of items tapping a particular concept or variable (as per the following example) are then summated for every respondent. This is an interval scale and the differences in the responses between any two points on the scale remain the same.

Example

Using the preceding Likert scale, state the extent to which you agree with each of the following statements:

My work is very interesting	1	2	3	4	5
I am not engrossed in my work all day	1	2	3	4	5
Life without my work will be dull	1	2	3	4	5

d. Semantic Differential Scale

Several bipolar attributes are identified at the extremes of the scale, and respon- dents are asked to indicate their attitudes, on what may be called a semantic space, toward a particular individual, object, or event on each of the attributes. The bipolar adjectives used, for instance, would employ such terms a Good–Bad; Strong–Weak; Hot–Cold. The semantic differential scale is used t assess respondents' attitudes toward a particular brand, advertisement, object, o individual. The responses can be plotted to obtain a good idea of their percep tions. This is treated as an interval scale. An example of the semantic differentia scale follows.

Example

Responsive	— — — — — — —	Unresponsive
Beautiful	— — — — — — —	Ugly
Courageous	— — — — — — —	Timid

e. Numerical Scale

The numerical scale is similar to the semantic differential scale, with the difference that numbers on a 5-point or 7-point scale are provided, with bipolar adjectives at both ends, as illustrated below. This is also an interval scale.

Example: *How pleased are you with your new real estate agent?*

Extremely Pleased	7	6	5	4	3	2	1	Extremely Displeased

f. Itemized Rating Scale

A 5-point or 7-point scale with anchors, as needed, is provided for each item and the respondent states the appropriate number on the side of each item, or cir cles the relevant number against each item, as per the examples that follow. Th responses to the items are then summated. This uses an interval scale.

Example (i) *Respond to each item using the scale below, and indicate your response number on the line by each item.*

1	2	3	4	5
Very Unlikely	Unlikely	Neither Unlikely Nor Likely	Likely	Very Likely

1. I will be changing my job within the next 12 months. —
2. I will take on new assignments in the near future. —
3. It is possible that I will be out of this organization within the next 12 months. —

Note that the above is a *balanced rating scale* with a *neutral* point.

	Example: (ii) Circle the number that is closest to how you feel for the item below.

Not at All Interested	Somewhat Interested	Moderately Interested	Very Much Interested
1	2	3	4

How would you rate your interest in changing current organizational policies?	1	2	3	4

This is an unbalanced rating scale which does not have a neutral point.

g. Fixed or Constant Sum Scale

The respondents are here asked to distribute a given number of points across various items as per the example below. This is more in the nature of an ordinal scale.

Example

In choosing a toilet soap, indicate the importance you attach to each of the following five aspects by allotting points for each to total 100 in all.

Fragrance —

Color —

Shape —

Size —

Texture of lather —

Total points 100

h. Stapel Scale

This scale simultaneously measures both the direction and intensity of the atti tude toward the items under study. The characteristic of interest to the study i placed at the center and a numerical scale ranging, say, from + 3 to – 3, on either side of the item as illustrated below. This gives an idea of how close or distant the individual response to the stimulus is, as shown in the example below. Since this does not have an absolute zero point, this is an interval scale.

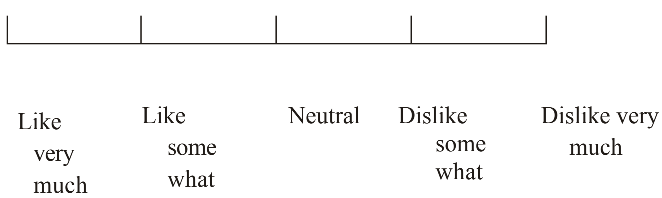
State how you would rate your supervisor's abilities with respect to each of the characteristics mentioned below, by circling the appropriate number.

+3	+3	+3
+2	+2	+2
+1	+1	+1
Adopting Modern Technology	Product Innovation	Interpersonal Skills
–1	–1	–1
–2	–2	–2
–3	–3	–3

i. The graphic rating scale

It is quite simple and is commonly used in practice. Under it the various points are usually put along the line to form a continuum and the rater indicates his rating by simply making a mark (such as ✓) at the appropriate point on a line that runs from one extreme to the other. Scale-points with brief descriptions may be indicated along the line, their function being to assist the rater in performing his job. The following is an example of five-points graphic rating scale when we wish to ascertain people's liking or disliking any product:

How do you like the product? (Please check)



j. Consensus Scale

Scales are also developed by consensus, where a panel of judges selects certain items, which in its view measure the relevant concept. The items are chosen particularly based on their pertinence or relevance to the concept. Such a consensus scale is developed after the selected items are examined and tested for their validity and reliability. One such consensus scale is the Thurstone Equal Appearing Interval Scale, where a concept is measured by a complex process followed by a panel of judges. Using a pile of cards containing several descrip- tions of the concept, a panel of judges offers inputs to indicate how close or not the statements are to the concept under study. The scale is then developed based on the consensus reached. However, this scale is rarely used for measuring orga- nizational concepts because of the time necessary to develop it.

(ii) Ranking scales:

Under ranking scales (or comparative scales) we make relative judgements against other similar objects. The respondents under this method directly compare two or more objects and make choices among them. As already mentioned, ranking scales are used to tap preferences between two or among more objects or items (ordinal in nature). However, such ranking may not give definitive clues to some of the answers sought. For instance, let us say there are four product lines and the manager seeks information that would help decide which product line should get the most attention. Let us also assume that 35% of the respondents choose the first product, 25% the second, and 20% choose each of products three and four as of importance to them. The manager cannot then conclude that the first product is the most preferred since 65% of the respondents did not choose that product! Alternative methods used are the *paired comparisons, forced choice,* and the *comparative* scale, which are dis- cussed below. There are three generally used approaches of ranking scales viz.

a. Paired Comparison

The paired comparison scale is used when, among a small number of objects, respondents are asked to choose between two objects at a time. This helps to assess preferences. If, for instance, in the previous example, during the paired comparisons, respondents consistently show a preference for product one over products two, three, and four, the manager reliably understands which product line demands his utmost attention. However, as the number of objects to be com- pared increases, so does the number of paired comparisons. The paired choices for *n* objects will be [(*n*) (*n*–1)/2]. The greater the number of objects or stimuli, the greater the number of paired comparisons presented to the respondents, and the greater the respondent fatigue. Hence paired comparison is a good method if the number of stimuli presented is small.

b. Forced Choice

The forced choice enables respondents to rank objects relative to one another among the alternatives provided. This is easier for the respondents, particularly if the number of choices to be ranked is limited in number.

Rank the following magazines that you would like to subscribe to in the order of preference, assigning 1 for the most preferred choice and 5 for the least preferred.

Fortune	—
India Today	—
Time	—
People	—
Nature	—

c. Method of rank order

Under this method of comparative scaling, the respondents are asked to rank their choices. This method is easier and faster than the method of paired comparisons stated above. For example, with 10 items it takes 45 pair comparisons to complete the task, whereas the method of rank order simply requires ranking of 10 items only. The problem of transitivity (such as *A* prefers to *B, B* to *C*, but *C* prefers to *A*) is also not there in case we adopt method of rank order. Moreover, a complete ranking at times is not needed in which case the respondents may be asked to rank only their first, say, four choices while the number of overall items involved may be more than four, say, it may be 15 or 20 or more. To secure a simple ranking of all items involved we simply total rank values received by each item. There are methods through which we can as well develop an interval scale of these data. But then there are limitations of this method. The first one is that data obtained through this method are ordinal data and hence rank ordering is an ordinal scale with all its limitations.

Then there may be the problem of respondents becoming careless in assigning ranks particularly when there are many (usually more than 10) items.

5. Criteria for the Good Test

There are two criteria to decide whether the scale selected is good or not. They are:

a. Reliability Test; and

b. Validity Test

a. Reliability Test

Reliability means the extent to which the measurement process is free from errors. Reliability deals with accuracy and consistency. The scale is said to be reliable, if it yields the same results when repeated measurements are made under constant conditions. *Reliability* refers to the extent to which a scale produces consistent results if repeated measurements are made. Systematic sources of error do not have an adverse impact on reliability, because they affect the measurement in a constant way and do not lead to inconsistency. In contrast, random error produces inconsistency, leading to lower reliability. Reliability is assessed by determining the proportion of systematic variation in a scale. This is done by determining the association between scores obtained from different administrations of the scale. If the association is high, the scale yields consistent results and is therefore reliable. Approaches for assessing reliability include the test–retest, alternative-forms and internal consistency methods.

Test–retest reliability

In **test–retest reliability**, respondents are administered identical sets of scale items at two different times, under as nearly equivalent conditions as possible. The time interval between tests or administrations is typically two to four weeks. The degree of similarity between the two measurements is determined by computing a correlation coefficient. The higher the correlation coefficient, the greater the reliability.

Alternative-forms reliability

In **alternative-forms reliability**, two equivalent forms of the scale are constructed. The same respondents are measured at two different times, usually two to four weeks apart (e.g. by initially using Likert scaled items and then using Stapel scaled items). The scores from the administrations of the alternative scale forms are correlated to assess reliability.

Internal consistency reliability

Internal consistency reliability is used to assess the reliability of a summated scale where several items are summed to form a total score. In a scale of this type, each item measures some aspect of the construct measured by the entire scale, and the items should be consistent in what they indicate about the construct. This measure of reliability focuses on the internal consistency of the set of items forming the scale.

Split-half reliability

The simplest measure of internal consistency is **split-half reliability**. The items on the scale are divided into two halves and the resulting half scores are correlated. High correlations between the halves indicate high internal consistency. The scale items can be split into halves based on odd- and even-numbered items or randomly. The problem is that the results will depend on how the scale items are split.

Coefficient alpha, or Cronbach's alpha

A popular approach to overcoming this problem is to use the coefficient alpha. The **coefficient alpha**, or Cronbach's alpha, is the average of all possible split-half coefficients resulting from different ways of splitting the scale items. This coefficient varies from 0 to 1, and a value of 0.6 or less generally indicates unsatisfactory internal consistency reliability. An important property of coefficient alpha is that its value tends to increase with an increase in the number of scale items. Therefore, coefficient alpha may be artificially, and inappropriately, inflated by including several redundant scale items.

Another coefficient that can be employed in conjunction with coefficient alpha is coefficient beta. Coefficient beta assists in determining whether the averaging process used in calculating coefficient alpha is masking any inconsistent items.

Example:

<u>Attitude towards a product or brand preference.</u>

Reliability can be ensured by using the same scale on the same set of respondents, using the same method. However, in actual practice, this becomes difficult as:

(i) Extent to which a scale produces consistent results

(ii) Test-retest Reliability: Respondents are administered scales at 2 different times under nearly equivalent conditions

(iii) Alternative-form Reliability: 2 equivalent forms of a scale are constructed, then tested with the same respondents at 2 different times

(iv) Internal Consistency Reliability:

(a) The consistency with which each item represents the construct of interest

(b) Used to assess the reliability of a summated scale

(c) Split-half Reliability

(v) Items constituting the scale divided into 2 halves, and resulting half scores are correlated: Coefficient alpha (most common test of reliability)

(vi) Average of all possible split-half coefficients resulting from different splitting of the scale items.

a. Validity Test

The paradigm of validity focused in the question "Are we measuring, what we think, we are measuring?" Success of the scale lies in measuring "What is intended to be measured?" Of the two attributes of scaling, validity is the most important.

*The **validity** of a scale may be considered as the extent to which differences in observed scale scores reflect true differences among objects on the characteristic being measured, rather than systematic or random error.*

There are several methods to check the validity of the scale used for measurement:

Construct Validity: A sales manager believes that there is a clear relation between job satisfaction for a person and the degree to which a person is an extrovert and the work performance of his sales force. Therefore, those who enjoy high job satisfaction, and have extrovert personalities should exhibit high performance. If they do not, then we can question the construct validity of the measure.

Content Validity: A researcher should define the problem clearly. Identify the item to be measured. Evolve a suitable scale for this purpose. Despite these, the scale may be criticised for being lacking in content validity. Content validity is known as face validity. An example can be the introduction of new packaged food. When new packaged food is introduced, the product representing a major change in taste. Thousands of consumers may be asked to taste the new packaged food. Overwhelmingly, people may say that they liked the new flavour. With such a favourable reaction, the product when introduced on a commercial scale may still meet with failure. So, what is wrong? Perhaps a crucial question that was omitted. The people may be asked if liked the new packaged food, to which the majority might have "yes" but the same respondents were not asked, "Are you willing to give up the product which you are consuming currently?" In this case, the problem was not clearly identified and the item to be 'measured' was left out.

Predictive Validity: This pertains to "How best a researcher can guess the future performance from the knowledge of attitude score"?

6. Questionnaire construction

Questionnaire is a structured technique for data collection consisting of a series of questions, written or verbal, that a respondent answers. A **questionnaire**, whether it is called a schedule, interview form or measuring instrument, is a formalised set of questions for obtaining information from respondents. Typically, a questionnaire is only one element of a data collection package that might also include (1) fieldwork procedures, such as instructions for selecting, approaching and questioning respondents: (2) some reward, gift or payment offered to respondents; and (3) communication aids, such as maps, pictures, advertisements and products (as in personal interviews) and return envelopes (in mail surveys). Any questionnaire has three specific objectives.

First, it must translate the information needed into a set of specific questions that the respondents can and will answer. Developing questions that respondents can and will answer and that will yield the desired information is difficult. Two apparently similar ways of posing a question may yield different information. Hence, this objective is most challenging.

Second, a questionnaire must uplift, motivate and encourage the respondent to become involved, to cooperate, and to complete the task. Figure 13.1 uses a basic marketing model of exchange of values between two parties to illustrate this point. Before designing any questionnaire or indeed any research technique, the researcher must evaluate 'what is the respondent going to get out of this?' In other words, the marketing researcher must have an empathy with target respondents and appreciate what they think when approached and questioned. Such an appreciation of what respondents go through affects the design of how they are approached, the stated purpose of the research, the rewards for taking part and the whole process of questioning and question design.

Third, a questionnaire should minimise response error. The potential sources of error in research designs were discussed in Chapter 3, where response error was defined as the error that arises when respondents give inaccurate answers or when their answers are misrecorded or mis-analysed. A questionnaire can be a major source of response error. Minimising this error is an important objective of questionnaire design.

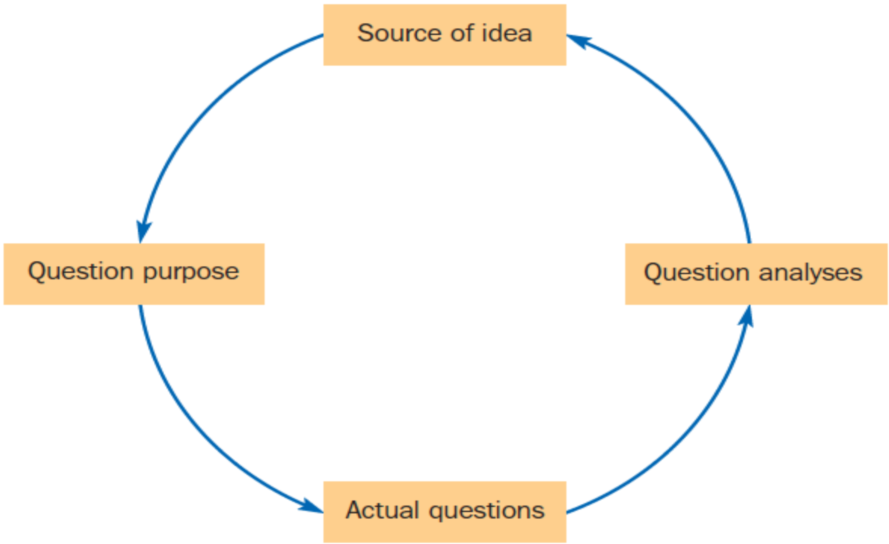
7. Questionnaire Design Process

The design process is founded upon generating information that will effectively support decision-makers. Establishing the nature of social problems and corresponding research problems, i.e. defining the nature of effective support. Different techniques and sources of information were outlined to help in the diagnosis process, which feed directly into the stages set out below:

a. The **'source of idea'** represents the culmination of marketing decision-maker and marketing researchers' diagnoses and the information they have available at the time of commissioning a marketing research project.
b. From the diagnoses, and the statement of marketing and research problems, emerge specific research questions. Based upon the diagnoses, the purpose of each potential question should be established, i.e. '**question purposes**'. Some research problems may be tackled through actual measurements in questionnaires. Other research problems may not be tackled by questionnaires.
c. With clear question purposes, the process of establishing **'actual questions'** can begin. At this point, the researchers have to put themselves 'in the shoes' of the potential respondent. It is fine to say that certain questions need to be answered, but this has to be balanced with an appreciation of whether respondents are able or indeed willing to answer particular questions, as illustrated in the following example.
d. Deciding how the data collected are to be analysed does not happen when questionnaires have been returned from respondents. **'Question analyses'** must be thought through from an early stage. The connections between questions and the

appropriate statistical tests that fulfil the question purposes should be established as the questionnaire is designed. Again, trade-offs have to be considered. In Chapter 12, different scale types were linked to different statistical tests. As one progresses from nominal through ordinal to interval and then ratio scales, more sophisticated statistical analyses can be performed. However, as one progresses through these scale types, the task for respondents becomes more onerous.

e. The understanding that is taken from the data comes back to the 'source of idea'. By now the researcher or questionnaire designers may have collected other data, interpreted existing data differently, or been exposed to new forces in the marketplace. They may even now see what questions they should have been asking!



8. Types of Questionnaire

A question may be unstructured or structured. We define unstructured questions and discuss their relative advantages and disadvantages and then consider the major types of structured questions: multiple choice, dichotomous and scales.

a. Unstructured questions

Unstructured questions are open-ended questions that respondents answer in their own words. They are also referred to as free-response or free-answer questions. The following are some examples:

- What is your occupation?
- What do you think of people who patronise secondhand clothes shops?
- Who is your favourite film personality?

Open-ended questions can be good first questions on a topic. They enable the respondents to express general attitudes and opinions that can help the researcher interpret their responses to structured questions. They can also be useful as a final question in a questionnaire.

After respondents have thought through and given all their answers in a questionnaire, there may be other issues that are important to them and that may not have been covered. Having an open-ended question at the end allows respondents to express these issues. As well as providing material to help the researcher interpret other responses, the respondents have the chance to express what they feel to be important Unstructured questions have a much less biasing influence on response than structured questions. Respondents are free to express any views. Their comments and explanations can provide the researcher with rich insights.

A principal disadvantage is that potential for interviewer bias is high. Whether the interviewers record the answers verbatim or write down only the main points, the data depend on the skills of the interviewers. Recorders should be used if verbatim reporting is important.

b. Structured questions

Structured questions specify the set of response alternatives and the response format. A structured question may be multiple choice, dichotomous or a scale.

i. Multiple-choice questions

In multiple-choice questions, the researcher provides a choice of answers and respondents are asked to select one or more of the alternatives given. Consider the following question:

Do you intend to buy a new car within the next six months?

______________________________ Definitely will not buy
______________________________ Probably will not buy
______________________________ Undecided
______________________________ Probably will buy
______________________________ Definitely will buy
______________________________ Other (please specify)

Of concern in designing multiple-choice questions are the number of alternatives that should be included and the order of potential responses, known as position bias. The response alternatives should include the set of all possible choices. The general guideline is to list all alternatives that may be of importance and to include an alternative labelled 'other (please specify)', as shown above. The response alternatives should be mutually exclusive.

Respondents should also be able to identify one, and only one, alternative, unless the researcher specifically allows two or more choices (e.g. 'Please indicate all the brands of soft drinks that you have consumed in the past week'). If the response alternatives are numerous, consider using more than one question to reduce the information processing demands on the respondents.

ii. Dichotomous questions

A dichotomous question has only two response alternatives, such as yes or no, or agree or disagree. Often, the two alternatives of interest are supplemented by a neutral alternative, such as 'no opinion', 'don't know', 'both' or 'none', as in this example. The question asked before about intentions to buy a new car as a multiple-choice question can also be asked as a dichotomous question.

Do you intend to buy a new car within the next six months?

✓

Yes

No

Don't know

iii. Scales

To illustrate the difference between scales and other kinds of structural questions, consider the question about intentions to buy a new laptop computer. One way of framing this using a scale is as follows:

Do you intend to buy a new car within the next six months?

Definitely will not buy	Probably will not buy	Undecided	Probably will buy	Definitely will buy
1	2	3	4	5

Unit 5: Sampling Design and Data Collection

(Dr. Sajal Das)

Scaling Techniques of Data Measurement:

The 4 types of scaling techniques for any Data are

(i) Nominal Data – Example Class Roll Number, Player's Jersey Number
(ii) Ordinal Data – Example Low Income Group, Middle Income Group, Higher Income Group
(iii) Interval Scale – Example Temperature measured in degree centigrade or Fahrenheit
(iv) Ratio Scale – is the highest level of scale and possesses all the properties of the earlier 3 scales.

In the popular SPSS Statistics software both interval and ratio are clubbed together as Scale, so at the time of data entry in the software we will see only 3 types: Nominal, Ordinal and Scale. Moreover, Nominal Data and Ordinal Data is known as Non-metric Data whereas Interval Scale and Ratio Scale together known as Metric Scale.

Sources of Data:

(i) Primary Sources: Originally collected for present investigation like through interviews, questionnaires, email surveys and the like.
(ii) Secondary Sources: Company's Balance Sheet are primary data for the company but secondary data for the stock market analyst.

Advantages of Secondary Data:

(i) Economical in terms of cost, time and effort.
(ii) In many complicated, sensitive and large sized enquires primary data cannot be collected.
(iii) One may not have sufficient organisational structure to collect primary data in all situation.

Organisation of Data:

(i) Construction of frequency table
(ii) Construction of Contingency table
(iii) Graphic or diagrammatic presentation

Sampling Techniques:

A sample is a subset of the population. It is quite often used to draw inferences about the population.

1. Kolkata traffic police wants to find out the pollution emission from trucks entering Kolkata city. A sample of 160 trucks may be taken for this purpose.

Sampling Techniques are broadly classified as Probability Sampling techniques and Non-probability sampling techniques. Probability sampling techniques are also known as statistical or random sampling techniques. In probability sampling, the probability of selection is known.

(i) Simple Random Sampling
(ii) Lottery Method

(iii) Random Numbers
(iv) Stratified Random Sampling
(v) Systematic Sampling
(vi) Cluster Sampling

Whereas for Non-probability sampling techniques:

(i) Judgmental or Purposive Sampling
(ii) Quota Sampling
(iii) Convenience Sampling.

Errors in Sampling: whether data are collected through census or sampling there is no guarantee that they will be free from error. Since samples do not include all the items there are chances that conclusions based on them may have errors. All errors in sampling may be divided into 2 categories:

(i) Sampling errors: It refers to the difference between parameter and statistic. For example, mean age of residents of small town may be 34 years whereas the mean age of a sample drawn from that small town of 100 families may be 26 years. Thus, there is a sampling error of -8 years in the estimation of mean age of the population. This extent of deviation of a particular statistic from the parameter is also referred to as precision. Thus, we can calculate sampling error under random sampling. The main reason of sampling error is that a sample may not have all the characteristics of the population.
(ii) Non-sampling errors are those which are related to data collection, measurement and recording. Even response errors like do not know, do not wish to answer or deliberately providing inaccurate answer are included in non-sampling error category.

Statistic and Parameter:

Let us now understand the basic difference between two key terms which are often mistakenly used interchangeably. Characteristics of the population like mean, median, proportion, standard deviation etc. which describe the entire population data set, are called Parameters. Parameters are calculated from population data and expressed in numerical value.

Characteristics of the sample expressed in numerical value are called statistic. They are calculated from sample data. Statistic is often used to estimate parameter. As a matter of convention, parameters are denoted by Greek letters or Capital letters and the lower case or Latin letters designate statistic.

By using sampling techniques, researches compute statistics and then employ them to estimate population.

Module II

Unit 7: Univariate Data Analysis

(Dr. Bikram Singh)

i. Univariate Data Analysis Meaning

Univariate analyses are used extensively in quality of life research. Univariate analysis is defined as analysis carried out on only one ("uni") variable ("variate") to summarize or describe the variable. Univariate analyses are contrasted with bivariate analyses (analyses involving two variables) and multivariate analyses (analyses of two or more variables simultaneously). There are three major characteristics of a single variable that we tend to look at namely, the distribution, the central tendency and the dispersion

ii. Descriptive Statistics Meaning

Descriptive statistics are numbers that are used to summarize and describe data. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data. Several descriptive statistics are often used at one time to give a full picture of the data. Descriptive statistics are just descriptive. Descriptive statistics are typically distinguished from inferential statistics. In descriptive statistics the objective is to describe what is or what the data shows. With inferential statistics, the objective is to try to reach conclusions that extend beyond the immediate data alone.

iii. Measures of Central Tendency

Measures of central tendency are the most basic and, often, the most informative description of a population's characteristics. They describe the "average" member of the population of interest. There are three measures of central tendency:

- ***Mean:*** It is the sum of a variable's values divided by the total number of values.
- ***Median***: the middle value of a variable. Median divides an arrayed series in two equal parts. Just as we have one median dividing a series in two equal parts, so three items would divide it in four parts (quartiles), nine items in ten parts (deciles) and ninety nine items in hundred parts (percentiles). Median is the best average in open and grouped distributions.
- ***Mode***: the value of the variate which occurs most frequently. It represents the most frequent value of a series. Mode is particularly useful average for discrete series and for a very large frequency.

The mean is the most commonly used measure of central tendency. Medians are generally used when a few values are extremely different from the rest of the values (this is called a skewed distribution).

iv. Empirical Relationship between Mean, Median and Mode

A distribution where the values of the Mean, Median and Mode coincide is called a symmetrical and if the above values are not equal then the distribution is said to be asymmetrical or skewed. In a moderately skewed distribution, there is a relation amongst Mean, Median and Mode. This is as follows:

Mean- Mode = 3(Mean –Median)

v. **Measures of Dispersion**

Measures of dispersion provide information about the spread of a variable's values. A measure of dispersion is designed to state the extent to which individual observations (or items) vary from their average.

Measures of dispersion are mainly of two types:

a. Absolute measures
    - Range
    - Quartile deviation
    - Mean Deviation
    - Standard Deviation
b. Relative Measures
    - Coefficient of Range
    - Coefficient of Quartile Deviation
    - Coefficient of Dispersion
    - Coefficient of Variation

Range is simply the difference between the smallest and largest values in the data Max Value- Min Value). For grouped frequency distribution range is the difference between the upper limit of the highest class interval and the lower limit of the lowest class interval. The interquartile range is the difference between the values at the 75^{th} percentile and the 25^{th} percentile of the data. ***Coefficient of Range*** = Range/(Max value +Min Value)

Quartile Deviation is half of the difference between the upper and lower quartiles i.e., ½ (Q_3 –Q_1). By dividing the quartile deviation by the average value of the two quartiles we will get ***Coefficient of Quartile Deviation***.

Mean Deviation of a series is the arithmetic average of the deviations of various items from the median or mean of that series. ***Coefficient of mean dispersion (about mean)*** **=** mean deviation/arithmetic mean. ***Coefficient of mean dispersion (about median)*** **=** mean deviation/median.

Variance is the most commonly used measure of dispersion. It is calculated by taking the average of the squared differences between each value and the mean.

Standard deviation, another commonly used statistic, is the square root of the variance. The Standard Deviation is a more accurate and detailed estimate of dispersion because an outlier can greatly exaggerate the range.

Coefficient of Variation is the ratio of the standard deviation to the mean expressed as a percentage. According to Karl Pearson, coefficient of variation is the percentage variation in the mean, while standard deviation is the total variation in the mean.

Skewness is a measure of whether some values of a variable are extremely different from the majority of the values. A variable is positively skewed if the extreme values are higher than the majority of values. A variable is negatively skewed if the extreme values are lower than the majority of values.

Analysis of Variance

Analysis of variance (ANOVA) is an analysis tool used in statistics that splits an observed aggregate variability found inside a data set into two parts: systematic factors and random factors. The systematic factors have a statistical influence on the given data set, while the random factors do not. The method is based upon an unusual result that the equality of several population means can be tested by comparing the sample variances using *F* distribution. In *t* statistic we test whether two population means are equal. The analysis of variance is an extension of the *t* test for the case of more than two means.

One-Way ANOVA Versus Two-Way ANOVA

There are two types of ANOVA: one-way (or unidirectional) and two-way. One-way or two-way refers to the number of independent variables in the analysis of variance test. A one-way ANOVA evaluates the impact of a sole factor on a sole response variable. It determines whether all the samples are the same. The one-way ANOVA is used to determine whether there are any statistically significant differences between the means of three or more independent (unrelated) groups.

A two-way ANOVA is an extension of the one-way ANOVA. With a one-way, you have one independent variable affecting a dependent variable. With a two-way ANOVA, there are two independents. For example, a two-way ANOVA allows a company to compare worker productivity based on two independent variables, such as salary and skill set. It is utilized to observe the interaction between the two factors and tests the effect of two factors at the same time.

Example of One Way ANOVA

Example 1: The following table shows the retail prices (Rs. per kg.) of a commodity in some shops selected at random in four cities:

A	B	C	D
34	29	27	34
37	33	29	36
32	30	31	38
33	34	28	35

Carry out the analysis of variance to test the significance of the differences between prices of the commodity in the four cities. [Given, $F_{0.05}$ = 3.49 for (3, 12) degrees of freedom]

Solution

Each observation is reduced by 39, and shown below:

Calculation for Analysis of Variance

	A	B	C	D	
	-5	-10	-12	-5	
	-2	-6	-10	-3	
	-7	-9	-8	-1	
	-6	-5	-11	-4	
Total	T_1= -20	T_2= -30	T_3= -41	T_4= -13	T = -104
Total of Squares	114	242	429	51	$\sum\sum x_{ij}^2$= 836
sample size	n_1= 4	n_2= 4	n_3= 4	n_4= 4	N= 16

Correction Factor (C.F.)= $T^2/N = (-104)^2/16 = 10816/16 = 676$

Total Sum of Squares (SS) = $\sum\sum xij^2$ - C.F. = 836- 676 = 160

Sum of Squares Between Groups (SSB) = $\sum(T_i^2/n_i)$ - C.F.

$= (-20^2/4 + -30^2/4 + -41^2/4 + -13^2/4) - 676$

= 787.50- 676

= 111.50

Sum of Squares due to Errors (SSE) = Total SS –SSB = 160-111.50 = 48.50

Analysis of Variance Table

Source of Variation	S.S	d.f	Mean Squares (M.S)	Observed F Value	Tabulated F Value
Between Groups	111.50	(k-1)= 4-1 = 3	111.50/3= 37.17	MSB/MSE= 37.17/4.04= 9.196	$F_{0.05}$ = 3.49 for (3,12) d.f
Within Groups (Errors)	48.50	(N-K)= 16-4= 12	48.50/12= 4.04		
Total	160	16-1 = 15			

Since the observed value of *F* (i.e., 9.196) exceeds the 5% tabulated value (i.e., 3.49) for (3,12) d.f., we reject the null hypothesis of equality of population means, and conclude that the retail prices of the commodity in the four cities are not equal.

In order to test which of the cities differ in prices, we calculate the critical difference (C.D.).

C.D. = $s\sqrt{2n}\ t_{0.025}$ (for 12d.f.)

s^2 = MSE

s = √MSE = √4.04

C.D. = √4.04* √2*4 * 2.18 = 12.39

The sample totals (of the reduced observations) are T_1= -20, T_2= -30, T_3= -41, T_4= -13

We have

$|T_1 - T_2| = 10$

$|T_1 - T_3| = 21$

$|T_1 - T_4| = 7$

$|T_2 - T_3| = 11$

$|T_2 - T_4| = 17$

$|T_3 - T_4| = 28$

Comparing these figures with the C.D. (i.e., 12.39) we find that the cities A and C, B and D, C and D differ in prices. Cities A and B and A and D may be taken to be having same prices.

Example of Two Way ANOVA

Example 2: The following table gives the estimates of acreage of cultivable land but not cultivated land out of 100 acres of total land, as obtained by three investigators in each of three districts. Perform an analysis of variance to test whether there are significant differences between investigators and districts. [Given $F_{0.05}$ = 6.94 for 12 d.f. (2, 4)]

Investigator	District		
	I	II	III
A	23	28	26
B	24	25	27
C	24	22	26

Solution

Each observation is reduced by 24, and shown below:

Calculation for Analysis of Variance

Investigator	District			Total
	I	II	III	
A	-1	4	2	5
B	0	1	3	4
C	0	-2	2	0
Total (T_i)	-1	3	7	T= 9
sample size	n_1= 3	n_2= 3	n_3= 3	N = 9

Total of the squares of all figures

$\sum\sum xij^2 = (-1)^2 + (4)^2 + (2)^2 + (0)^2 + (1)^2 + (3)^2 + (0)^2 + (-2)^2 + (2)^2 = 39$

Correction Factor (C.F.) = $T^2/N = (9)^2/9 = 81/9 = 9$

Total Sum of Squares (SS) = $\sum\sum xij^2$ - C.F. = 39-9 = 30

Sum of Squares (SS) between Investigators = $(T_1^2 + T_1^2 + T_1^2)/3$ – C.F

$= (5^2 + 4^2 + 0^2)/3 - 9$

$= 41/3 - 9$

$= 4.67$

Sum of Squares (SS) between Districts = $(T_1'^2 + T_1'^2 + T_1'^2)/3$ – C.F.

$= [(-1)^2 + (3)^2 + (7)^2]/3 - 9$

$= 59/3 - 9$

$= 10.67$

SS due to Error = Total SS – (SS between investigators) - (SS between districts)

= 30 – 4.67 – 10.67

= 14.66

Analysis of Variance Table

Source of Variation	S.S	d.f	Mean Squares (M.S)	Observed F Value	Tabulated F Value
Between Investigators	4.67	3-1 = 2	4.67/2= 2.34	2.34/3.67 = 0.64	$F_{0.05}$ = 6.94 for (2,4) d.f
Between Districts	10.67	3-1 = 2	10.67/2 = 5.34	5.34/3.67 = 1.46	$F_{0.05}$ = 6.94 for (2,4) d.f
Within Groups (Errors)	14.66	4	14.66/4 = 3.67		
Total	30	9-1 = 8			

Since the observed value of *F for experimenters* (i.e., 0.64) is less than the corresponding tabulated value (i.e., 6.94) for d.f. (2, 4), it is not significant at 5% level. We conclude that the mean acreage of cultivable land in the three districts as determined by the three investigators may not be different from one another, i.e., there are no significant differences between investigators.

Since the observed value of *F for districts* (i.e., 1.46) is less than the corresponding tabulated value (i.e., 6.94) for d.f (2,4), it is not significant at 5% level. We conclude that the estimates of acreage of cultivable land in the three districts may not be different from one another, i.e., there are no significant differences between districts.

Paper GE.404: Business Research Methods

Unit 8: Multivariate Data Analysis

[Prof. (Dr.) Ashish Kumar Sana]

1. Multivariate Analysis: Meaning

In simple terms, multivariate means multi-variables i.e. engagement of more than two variables at a time. Multivariate analysis deals with the statistical analysis of data collected on more than one dependent variable. In other words, multivariate analysis is an analysis of statistical technique that analyse the relationship between more than two variables which shows the effect of more than one variable on one variable that is independent variable.

Multivariate analysis refers to all statistical techniques that simultaneously analyse multiple measurements on individuals or objects under investigation. Any simultaneous analysis of more than two variables can be considered as multivariate analysis.

[Source: Hair, Black, Babin, Anderson and Tatham, (2006), Multivariate Data Analysis, Pearson, 6th Impression)

So, multivariate data analysis is a set of statistical models that examine patterns in multidimensional data by considering, at once, several data variables. It is an expansion of bivariate data analysis, which considers only two variables in its models. As multivariate models consider more variables, they can examine more complex phenomena and find data patterns that more accurately represent the real world.

Basic Concept

Variate: It is a linear combination of variables with empirically determined weights. The variables are specified by the researcher, whereas the weights are determined by the multivariate technique relevant to a specific objective. A variate of n weighted variables (X_1 to X_n) can be defined mathematically as -

$$\text{Variate value} = w_1X_1+w_2X_2+\ldots\ldots+w_nX_n$$

Where X_n is the observed variable and w_n is the weight which is determined by the multivariate technique.

2. Uses or Significance of Multivariate Data Analysis

Multivariate analysis helps the researcher, analyst, business organisation, government, industries in decision making for the future. The main objective of multivariate techniques is to analyse the market scenario which explain what customer's habit, preferences, choices, target market campaigns, new product or serving Design and Refinement existing product. Multivariate techniques makes possible to conduct theoretically significant research and to evaluate the effects naturally occurring parametric variation in the manner in which usually they occur.

Few decades ago, it was very difficult task for the analysts and researchers for collecting and analysing large data but nowadays it becomes easy because of availability of computer software, statistical software packages and internet. So, it reduces the human task.

3. Measurement Scales

Measurement scale is a technique to measure the variable or to identify the data in selection of the appropriate multivariate method of analysis. It is important in accurately representing the concept of interest and is instrumental in the selection of the appropriate multivariate method of analysis. Data can be classified into generally two categories-Metric and Nonmetric.

Data Type	Attributes	Example	Scale
Metric	Quantitative	Male or Female	Nominal or Ordinal
Nonmetric	Qualitative	Level of satisfaction	Interval and Ratio

3.1 Nonmetric Measurement Scales

It generally measures quality of data which is qualitative in nature. Nonmetric measurement can be made with either a Nominal or an Ordinal scale.

Nominal Scales: A nominal scale assigns numbers as way to label or identify objects or subjects which have no quantitative meaning beyond indicating the presence or absence of the attribute or characteristic under investigation. Example: Individual's gender, occupation.

Ordinal Scales: In ordinal scale, variables will be ordered or stratified in respect to the number of the attribute possessed. Every subject or object can be compared with another in terms of a 'greater than' or 'less than' relationship. It provides no measure of the actual amount or magnitude in absolute terms, only the order of the values.

3.2 Metric Measurement Scales

Metric data are used when subjects differ in amount or degree on a particular attribute such as level of satisfaction or commitment to a job. Metric measurement scales are (i) interval scales and (ii) ratio scales.

Interval Scales: Interval scale refers to the level of measurement in which the attributes composing variables are measured on specific numerical scores or values and there are equal distances between attributes. The distance between any two adjacent attributes is called an interval, and intervals are always equal.

[Source: Encyclopaedia of Research Design, Neil J Salkind, (2010)]

Ratio Scales: Ratio scale refers to the level of measurement in which the attributes composing variables are measured on specific numerical scores or values that have equal distances between attributes or points along the scale and are based on a "true zero" point. Among four levels of measurement, including nominal, ordinal, interval, and ratio scales, the ratio scale is the most precise.

[Source: Encyclopaedia of Research Design, Neil J Salkind, (2010)]

4. Multivariate Techniques

There are two categories of multivariate techniques, each pursuing a different type of relationship in the data: (i) dependence and (ii) interdependence.

Dependence relates to cause-effect situations and tries to see if one set of variables can describe or predict the values of other ones. **Interdependence** refers to structural inter-correlation and aims to understand the underlying patterns of the data.

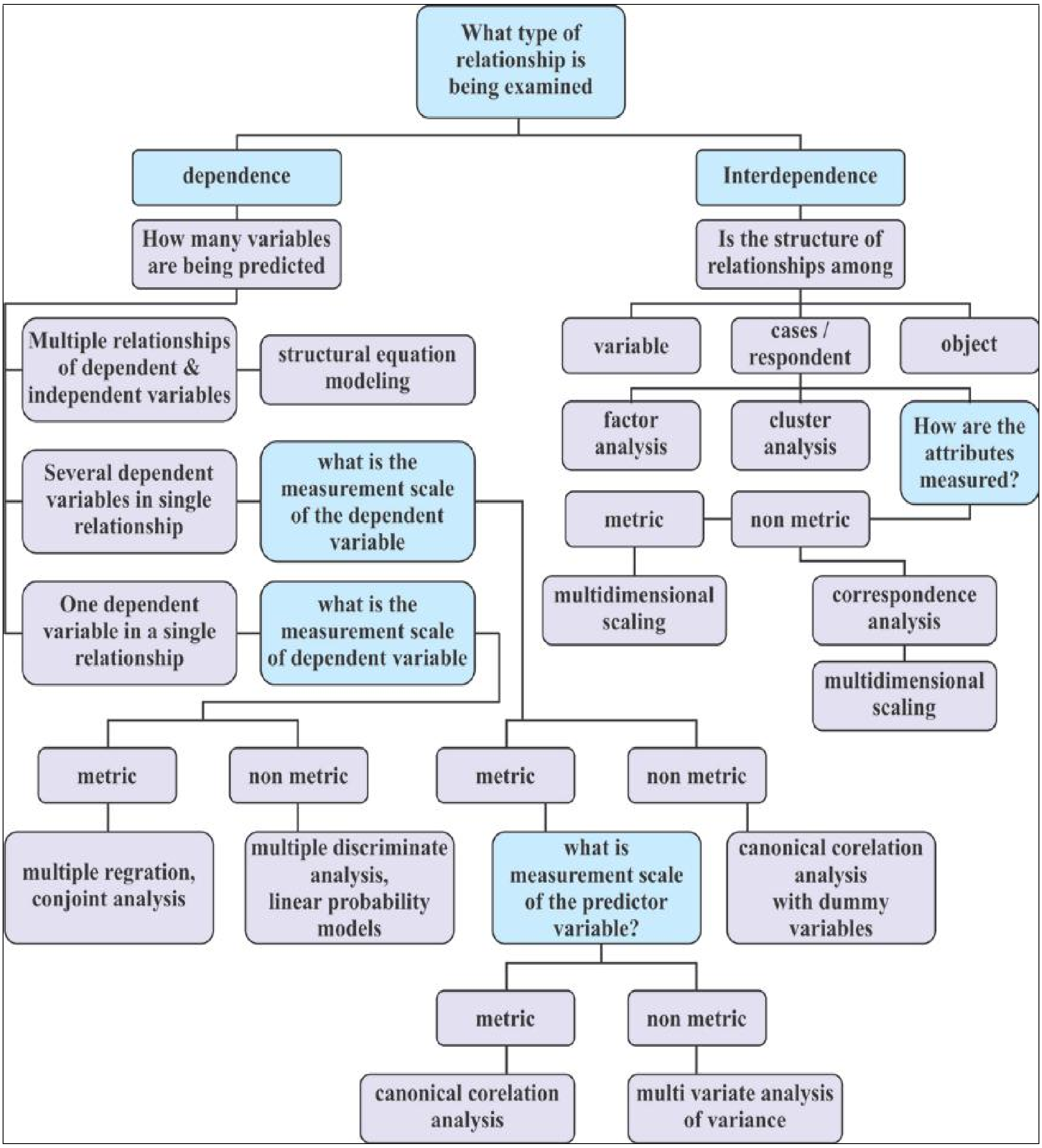
There are several multivariate models capable of finding those relationships, and many factors distinguish them. One of the primary factors that must be taken into account when choosing a technique is the nature of the data variables: they can be metric or non-metric.

Dependence Technique	Interdependence Technique

A **dependence technique** may be defined as one in which a variable or set of variables is identified as the dependent variable to be predicted or explained by other variables known as independent variables.

An example of a dependence technique is multiple regression analysis. In contrast, an interdependence technique is one in which no single variable or group of variables is defined as being independent or dependent. Rather, the procedure involves the simultaneous analysis of all variables in the set. Factor analysis is an example of an interdependence technique.

Multivariate Techniques: Figure 1



[Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson]

4.1 Dependence Techniques

The different dependence techniques can be categorized by two characteristics:

(1) The number of dependent variables and

(2) The type of measurement scale employed by the variables.

Dependence Techniques are:

1. Multiple Regression Analysis
2. Multiple Discriminant Analysis and Logistic Regression
3. Multivariate Analysis of Variance
4. Conjoint Analysis

Combinations	Appropriate Technique
Single Dependent Variable → Metric	Either multiple regression analysis or conjoint analysis*.
Single Dependent Variable → Nonmetric (categorical)	Multiple discriminant analysis and linear probability models.
Several Dependent Variables → Metric	We must then look to the independent variables ↓ If the independent variables are metric, canonical correlation is appropriate.
	If the independent variables are nonmetric, the technique of multivariate analysis of variance (MANOVA) should be selected.
Several Dependent Variables → Nonmetric	If the several dependent variables are non-metric, then they can be transformed through **dummy variable** coding (0–1) and canonical analysis can again be used.
Several Dependent/Independent relationships	Then structural equation modelling is appropriate.
Metric (Quantitative/Numerical) or Non-metric (Qualitative/Categorical) Dependent Variables.	To be specified later.

(Source: Multivariate Data Analysis, author: Hair | Black | Babin | Anderson | Tatham)

** Conjoint analysis is a special case. It involves a dependence procedure that may treat the dependent variable as either nonmetric or metric, depending on the type of data collected.*

4.2 Interdependence Techniques

Interdependence techniques are focused on the definition of structure, assessing interdependence without any associated dependence relationships. Interdependence techniques can be classified by variable, case or respondent and object wise.

Interdependence Techniques are:

1. Cluster Analysis
2. Multidimensional Scaling and Correspondence Analysis

5. Types of Multivariate Technique

The more established and emerging multivariate techniques are briefly discussed below:

5.1 Principal Components and Common Factor Analysis

Factor analysis can be used to examine the underlying patterns or relationships for a large number of variables and to determine whether the information can be condensed o summarised in a smaller set of factors or components. It detects groups of variables with high correlation, which the analyst may use as a basis to create a new variable that can replace them with little information loss. Factor analysis includes techniques such as principal component analysis and common factor analysis.

Component analysis is known as principal component analysis which considers the total variance and derives factors that contain small proportions of unique variance and in some instances error variances. Whereas, common factor analysis considers only the common or shared variance assuming that both the unique and error variance are not of interest in defining the structure of the variables.

[Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson]

5.2 Multiple Regression

Multiple regression analysis is a statistical technique that can be used to analyse the relationship between a single dependent variable and several independent variables. The objective of this technique is to predict the changes within the independent variables. The ever-widening applications of multiple regression fall into two broad classes of research problems: prediction and explanation. Prediction involves the extent to which the regression variate can predict the independent variable. Explanation examines the regression coefficients for each independent variable. These research problems are not mutually exclusive, rather it can address either or both types of research problem.

Multiple Regression Analysis		
Y_1	=	$X_1+X_2+X_3+\ldots\ldots\ldots+X_n$
(Metric/Quantitative)		(Metric, Non metric)
		(Quantitative, Qualitative)

[Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson]

5.3 Multiple Discriminant Analysis and Logistic Regression

Discriminant analysis and logistic regression are the appropriate statistical techniques when the dependent variable is a categorical (nonmetric) variable and the independent variables are metric variables. Discriminant analysis is capable of handling either two groups or multiple (three or more) groups. When two classifications are involved, the technique is called two-group Discriminant analysis. When there or more classifications are identified, the technique is referred as Multiple Discriminant Analysis (MDA).

Logistic regression is a specialised form of regression that is formulated to predict and explain a binary (two-group) categorical variable rather than a metric dependent measure. The form of the logistic regression variate is similar to the variate in multiple regression. The variate represents a single multivariate relationship with regression like coefficients indicating the relative impact of each independent variable.

Multiple Discriminate Analysis		
Y_1	=	$X_1+X_2+X_3+\ldots\ldots\ldots+X_n$
(Non metric/Qualitative)		(Metric/Quantitative)

[Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson]

5.4 Multivariate Analysis of Variance (MANOVA)

The multivariate analysis of variance (MANOVA) is a dependent technique that measures the differences for two or more metric dependent variables based on a set of categorical (nonmetric) variables acting as independent variables. MANOVA is an extension analysis of variance (ANOVA). ANOVA and MANOVA can be stated in the following general forms:

Analysis of Variance (ANOVA)

Y_1	=	$X_1+X_2+X_3+\ldots\ldots\ldots+X_n$
(Metric/Quantitative)	=	(Non metric/Qualitative)

Multivariate Analysis of Variance (MANOVA)

$Y_1+Y_2+Y_3+\ldots\ldots\ldots+Y_n$	=	$X_1+X_2+X_3+\ldots\ldots\ldots+X_n$
(Metric/Quantitative)		(Non metric/ Qualitative)

[Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson]

5.5 Conjoint Analysis

Conjoint analysis is a multivariate technique developed specially to understand how respondents develop preferences for any type of object (e.g. products, services, or ideas). Conjoint analysis can be used in almost any area in which decisions are studied. It assumes that any set of objects (e.g. brands, companies) or concepts (e.g. positioning, benefits, images) is evaluated as a bundle of attributes. After determination the contribution of each factor to the consumer's overall evaluation, the researcher can then:

(i) Define the object or concept with optimum combination of features,

(ii) Show the relative contributions of each attribute and each level to the overall evaluation of the object,

(iii) Use estimates of purchaser or customer judgements to predict preferences among objects with differing sets of features,

(iv) Isolate groups of potential customers that place differing importance on the features to define high and low potential segments,

(v) Identify marketing opportunities by exploring the market potential for combinations of features

Knowledge of the preference structure for each individual enables almost unlimited flexibility to examine both individual and aggregate reactions to a wide range of product or service related issues.

Conjoint analysis can be expressed as:

Conjoint Analysis

Y_1	=	$X_1+X_2+X_3+\ldots\ldots\ldots\ldots+X_n$
(Metric, Non metric)		
(Quantitative, Qualitative)		(Non metric/Qualitative)

[Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson]

5.6 Cluster Analysis

Cluster analysis is a group of multivariate techniques whose primary objective is to group of objects based on the characteristics they possess. Cluster analysis classifies objects (e.g. respondents, products, or other entities) so that each object is similar to others in the cluster based on a set of selected characteristics. The resulting clusters of objects should exhibit high internal (within-cluster) homogeneity and high external (between-cluster) heterogeneity. Thus, if the classification is successful, the objects within clusters will be close together when plotted geometrically and different clusters will be far apart.

[Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson]

5.7 Multidimensional Scaling or Perceptual Mapping and Correspondence Analysis

Multidimensional scaling (MDS) also known as perceptual mapping is a procedure that enables a researcher to determine the perceived relative image of a set of objects (e.g. firms, products, ideas, or other items associated with commonly held perceptions). The application MDS is appropriate when the objective is more oriented toward understanding overall preferences or perceptions rather than detailed perspectives involving individual attributes. To perform MDS analysis, the researcher performs three basic steps:

i) gathers measures of similarity or preference across the entire set of objects to be analysed,
ii) uses MDS techniques to estimate the relative position of each object in multidimensional space, and
iii) identifies and interprets the axes of the dimensional space in terms of perceptual and/or objective attributes.

Correspondence analysis (CA) is an interdependence technique that has become increasingly popular for dimensional reduction and perceptual mapping. Correspondence analysis provides a valuable analytical tool for a type of data (nonmetric) that often is not the focal point of multivariate techniques. It also provides the researcher with a complementary compositional technique to MDS for addressing issues where direct comparison of objects and attributes is preferable.

Correspondence analysis has three distinguishing characteristics:

(1) it is a compositional technique, rather than a decompositional approach, because the perceptual map is based on the association between objects and a set of descriptive

characteristics or attributes specified by the researcher;

(2) its most direct application is portraying the correspondence of categories of variables, particularly those measured in nominal measurement scales, which is then used as the basis for developing perceptual maps; and

(3) the unique benefits of CA lie in its abilities for representing rows and columns, for example, brands and attributes, in joint space.

(Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson

5.8 Structural Equation Modelling

Structural equation modelling examines the structure of interrelationships expressed in a series of equations, similar to a series of multiple regression equations. These equations depict all of the relationships among constructs (the dependent and independent variables) involved in the analysis. Structural equation modelling can be expressed as follows:

Structural Equation Modelling

$$Y_1 = X_{11} + X_{12} + X_{13} + \ldots\ldots\ldots + X_{1n}$$
$$Y_2 = X_{21} + X_{22} + X_{23} + \ldots\ldots\ldots + X_{2n}$$
$$Y_m = X_{m1} + X_{m2} + X_{m3} + \ldots\ldots\ldots + X_{mn}$$

(Metric/Quantitative) (Metric, Non metric)

(Quantitative, Qualitative)

(Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson

5.9 Canonical Correlation Analysis

Canonical correlation analysis is one of the most flexible dependence techniques. It supports many dependent variables and many independent, with any combinations of natures, meaning it is always an option on the table.

Canonical correlation analysis could be a logical extension of multiple correlation analysis. It involves single metric variable and several other metric independent variables. The target is to develop a linear combination of every set of variables to maximise the correlation between the two sets. The procedure involves getting a group of weights for the dependent and explanatory variables that offer the most straightforward correlation between the set of explanatory variables.

Assume a company conducts a study that collects information on its service quality based on answers to 50 metrically measured questions. The study uses questions from published service quality research and includes benchmarking information on perceptions of the service quality of “world-class companies” as well as the company for which the research is being conducted. Canonical correlation could be used to compare the perceptions of the world-class companies on the 50 questions with the perceptions of the company. The research could then conclude whether the perceptions of the company are correlated with those of world-class companies. The technique would provide information on the overall correlation of perceptions as well as the correlation between each of the 50 questions.

Canonical Correlation

$Y_1 + Y_2 + Y_3 + \ldots\ldots\ldots + Y_n = X_1 + X_2 + X_3 + \ldots\ldots\ldots + X_n$ (Metric, non metric)

(Metric, non metric) (Quantitative, Qualitative)

(Quantitative, Qualitative)

(Source: Hair, Joseph F., et al. Multivariate Data Analysis (2006) 9th Impression., Pearson

Suggested Readings:

- Hair, Joseph F., et al. *Multivariate Data Analysis (2006) 9th Impression., Pearson*
- Neil J Salkind, (2010), *Encyclopaedia of Research Design,* Sage Publications
- Alvin C. Rencher (2002), *Methods of Multivariate Analysis*, Second Edition, WILEY & SONS, INC

Paper GE.404:

Business Research Methods

Unit 8: Multivariate Data Analysis

Regression Analysis

[Dr. Swapan Sarkar]

Unit 8: THE NATURE OF REGRESSION ANALYSIS

HISTORICAL ORIGIN OF THE TERM REGRESSION

The term regression was introduced by Francis Galton. In a famous paper, Galton found that, although there was a tendency for tall parents to have tall children and for short parents to have short children, the average height of children born of parents of a given height tended to move or "regress" toward the average height in the population as a whole. In other words, the height of the children of unusually tall or unusually short parents tends to move toward the average height of the population. Galton's law of universal regression was confirmed by his friend Karl Pearson, who collected more than a thousand records of heights of members of family groups. He found that the average height of sons of a group of tall fathers was less than their fathers' height and the average height of sons of a group of short fathers was greater than their fathers' height, thus "regressing" tall and short sons alike toward the average height of all men. In the words of Galton, this was "regression to mediocrity."

THE MODERN INTERPRETATION OF REGRESSION

The modern interpretation of regression is, however, quite different. Broadly speaking, we may say:

Regression analysis is concerned with the study of the dependence of one variable, the dependent variable, on one or more other variables, the explanatory variables, with a view to estimating and/or predicting the (population) mean or average value of the former in terms of the known or fixed (in repeated sampling) values of the latter.

Examples:

1. Reconsider Galton's law of universal regression. Galton was interested in finding out why there was a stability in the distribution of heights in a population. But in the modern view our concern is not with this explanation but rather with finding out how the average height of sons' changes, given the fathers' height. In other words, our concern is with predicting the average height of sons knowing the height of their fathers.

2. Turning to economic examples, an economist may be interested in studying the dependence of personal consumption expenditure on after tax or disposable real personal income. Such an analysis may be helpful in estimating the marginal propensity to consume (MPC), that is, average change in consumption expenditure for, say, a dollar's worth of change in real income.

Note: In regression analysis we are concerned with what is known as the statistical, not functional or deterministic, dependence among variables, such as those of classical physics. In statistical relationships among variables we essentially deal with random or stochastic variables, that is, variables that have probability distributions. In functional or deterministic dependency, on the other hand, we also deal with variables, but these variables are not random or stochastic.

However, one point should be kept in mind that a statistical relationship in itself cannot logically imply causation. To ascribe causality, one must appeal to a priori or theoretical considerations. Thus, in the second example cited, one can invoke economic theory in saying that consumption expenditure depends on real income.

TERMINOLOGYAND NOTATION

Before we proceed to a formal analysis of regression theory, let us dwell briefly on the matter of terminology and notation. In the literature the terms dependent variable and explanatory variable are described variously. A representative list is:

Dependent variable	Explanatory variable
⇕	⇕
Explained variable	Independent variable
⇕	⇕
Predictand	Predictor
⇕	⇕
Regressand	**Regressor**
⇕	⇕
Response	Stimulus
⇕	⇕
Endogenous	Exogenous
⇕	⇕
Outcome	Covariate
⇕	⇕
Controlled variable	Control variable

If we are studying the dependence of a variable on only a single explanatory variable, such as that of consumption expenditure on real income, such a study is known as simple, or two-variable, regression analysis. However, if we are studying the dependence of one variable on more than one explanatory variable, as in the crop-yield, rainfall, temperature, sunshine, and fertilizer examples, it is known as multiple regression analysis. In other words, in two-variable regression there is only one explanatory variable, whereas in multiple regression there is more than one explanatory variable.

CONCEPT OF POPULATION REGRESSION FUNCTION

Each conditional mean $E(Y|Xi)$ is a function of Xi, where Xi is a given value of X. Symbolically,

$E(Y|Xi)= f(Xi)$ (2.2.1)

where f(Xi) denotes some function of the explanatory variable X. In our example, $E(Y|Xi)$ is a linear function of Xi. Equation (2.2.1) is known as the conditional expectation function (CEF) or population regression function (PRF) or population regression (PR) for short. It states merely that the expected value of the distribution of Y given Xi is functionally related to Xi. In simple terms, it tells how the mean or average response of Y varies with X.

As a first approximation or a working hypothesis, we may assume that the PRF $E(Y|Xi)$ is a linear function of Xi, say, of the type

$E(Y|Xi) = \beta_1 + \beta_2 Xi$

where β_1 and β_2 are unknown but fixed parameters known as the regression coefficients; β1 and β2 are also known as intercept and slope coefficients, respectively. This equation itself is known as the linear population regression function. Some alternative expressions used in the literature are linear population regression model or simply linear population regression. In the sequel, the terms regression, regression equation, and regression model will be used synonymously.

Department of Commerce
University of Calcutta

STUDY MATERIAL FOR INTERNAL CIRCULATION OF M.COM STUDENTS

M.Com/ 4th Semester

Subject: Business Research Methods
Paper: GE 404 / Module II

Name of the Faculty: Professor (Dr.) S.S.Saha

[DAY AND EVENING SHIFT]

TOPIC

Unit 9: Time Series Analysis

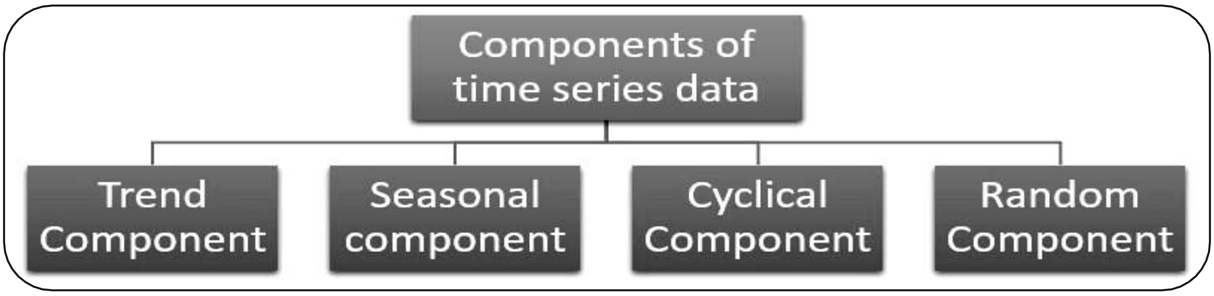
Time Series Analysis: Time series components; Box-Jenkins methodology; Evaluation of forecast accuracy; Unit root tests; Granger Causality, Co-integration and error correction mechanism; Volatility modeling ; GARCH models (basic idea only).

---

1. Time Series components

Component of Time Series Data

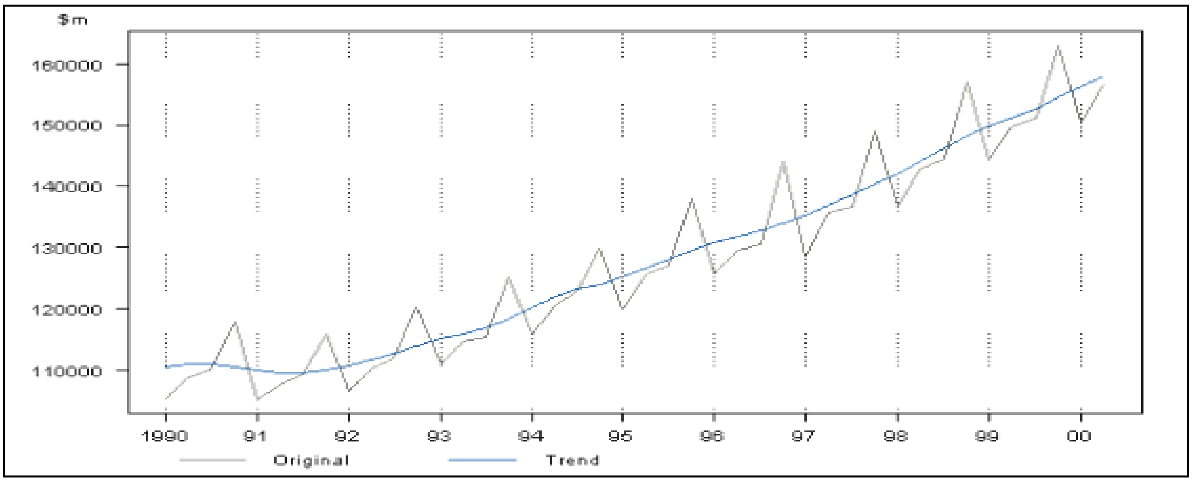
A time series is, a sequence of discrete-time data, a series of data points indexed in time order. A time series is considered as a collection of observations of well-defined data items obtained through repeated measurements over time. For instance, measuring the value of retail sales each month of the year would comprise a time series. The reason is that sales revenue is well defined, and consistently measured at equally spaced intervals. Data collected irregularly or only once are not time series. The component of time series data comprises four basic components: (a) Trend (secular trend or long term variation); (b) Seasonal effect (seasonal variation or seasonal fluctuations); (c) Other Cyclic Changes (cyclical variation or cyclic fluctuations); and (d) Other Irregular Variation (irregular fluctuations) or Random. These are discussed below:



(a) Secular Trend

The word trend means 'tendency'. So, secular trend is that component of the time series which gives the general tendency of the data for a long period. Trend is the long term movement in a time series without calendar related and irregular effects, and is a reflection of the underlying level. It is the result of influences such as population growth, price inflation and general economic changes. To realize the meaning of the long term, let for instance, climate variables sometimes show cyclic variation over a very long time period such as 60 years. If one just had 25 years of data, this long term fluctuation would appear to be a trend, but if several hundreds of years of data are available, long term fluctuations would be visible.

Figure 1: Quarterly Gross Domestic Product



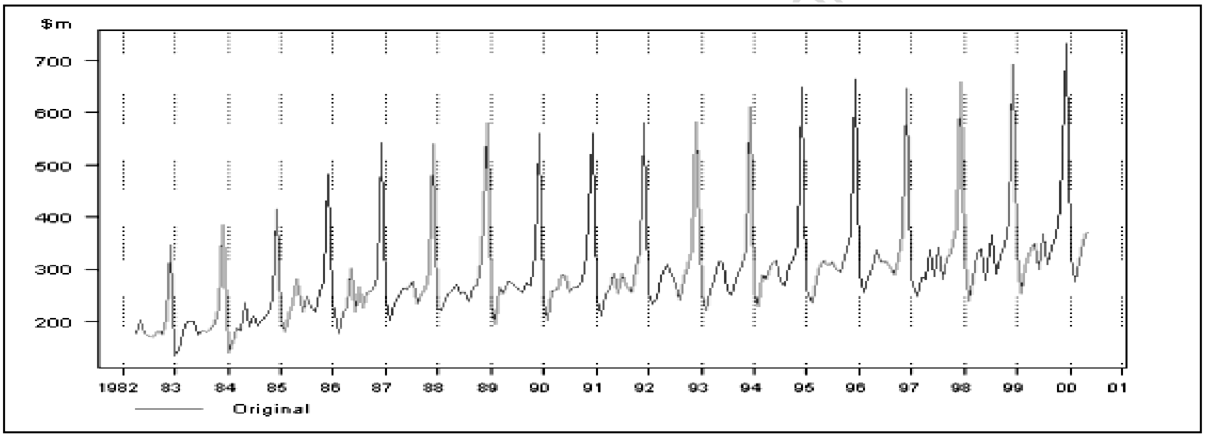
(Source : Australian Bureau of Statistics)

These movements are systematic in nature where the movements are broad, steady, showing a slow rise or fall in the same direction. The trend may be linear or non-linear (curvilinear). The secular trends examples are: increase in prices, increase in pollution, an increase in the need for wheat, increase in literacy rate, etc. Trend or long term variation is shown in the figure 1. The figure 1 portrays a series in which there is an obvious upward trend over time.

(b) Seasonal Effect

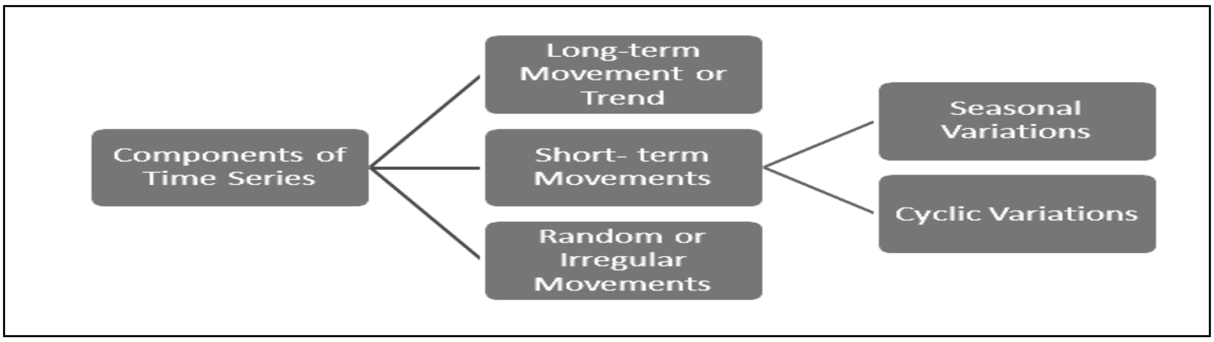
Seasonal fluctuations describe any regular variation with a period of less than one year. For instance, cost of various types of fruits and vegetables, clothes, unemployment figures, average daily rainfall, increase in the sale of tea in winter, increase in the sale of ice cream in summer, etc., all show seasonal variations. Basically, seasonal variations are caused by climate, social customs, religious activities, etc. However, a seasonal effect is a systematic and calendar related effect. For instance, the sharp rise in most retail series occurs around December in response to the Christmas period, or an increase in water consumption in summer due to warmer weather. Seasonal fluctuation is revealed in the figure 2. The figure shows a large seasonal increase in December retail sales in New South Wales due to Christmas shopping.

Figure 2: Monthly Retail Sales in New South Wales (NSW) Retail Department Stores



(Source : Australian Bureau of Statistics)

Seasonal and Cyclic Variations are the periodic changes or short-term fluctuations for more than a year.



(c) Other Cyclic Changes

Apart from seasonal variations, there is another type of fluctuation which usually lasts for more than a year. Time series shows cyclical variation at a fixed period due to some other physical cause, such as daily variation in temperature. Cyclical fluctuation, basically, is a non-seasonal component varying in a recognizable cycle. Sometimes series reveals variation which does not have a fixed period although it is predictable to some extent. For example, economic data is affected by business cycles with a period varying between about 5 and 7 years. In weekly or monthly data, the cyclical component may describe any regular variation in time series data. The cyclical fluctuation is periodic in nature and repeats itself like a business cycle, which has four phases: (a) Peak, (b) Recession, (c) Trough/Depression, and (d) Expansion. The cyclical variation is shown in the following figure.

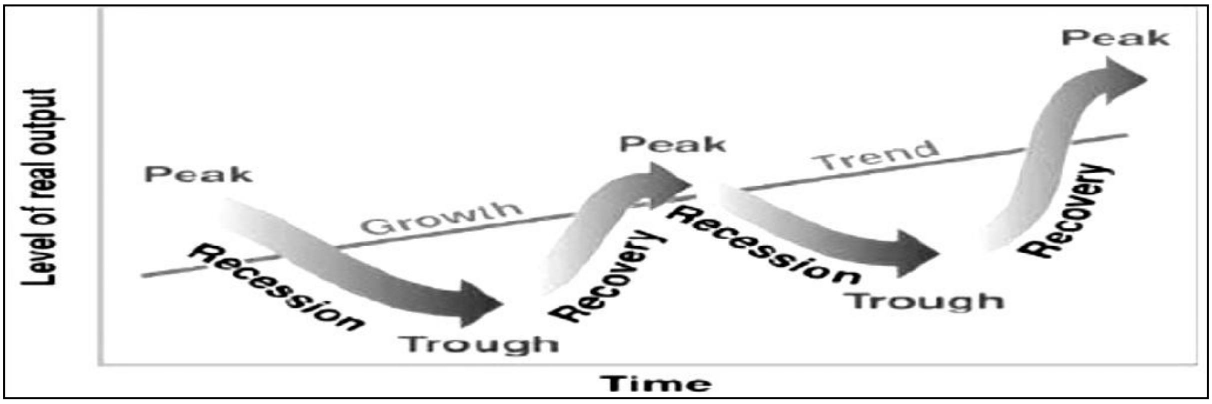
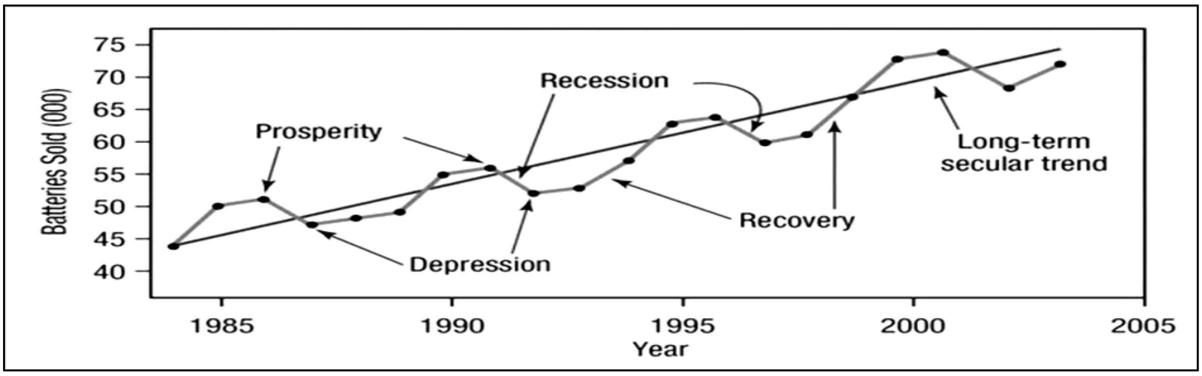


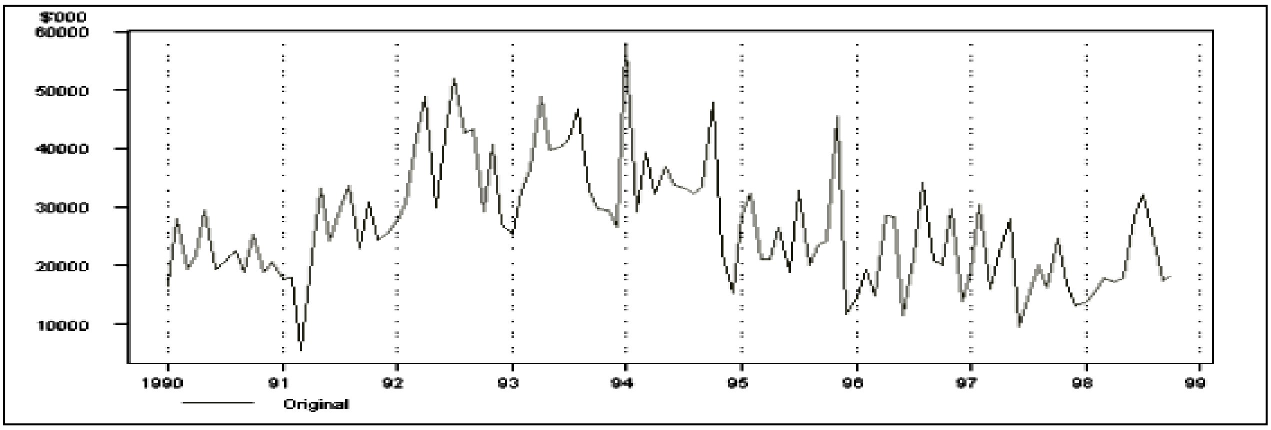
Figure 3: Batteries sold from 1984 to 2005



(d) Random / Irregular Fluctuations

The irregular component is also known as the residual. This is what remains after the seasonal and trend components of a time series have been estimated and removed. It results from short term variations in the series which are neither systematic nor predictable. In a highly irregular series, these fluctuations can dominate movements, which will cover the trend and seasonality. If the magnitude of the irregular component of a series is strong compared with the magnitude of the trend component, the underlying direction of the series can be distorted.

Figure 4: Monthly Value of Building Approvals, Australian Capital Territory (ACT)



(Source: Australian Bureau of Statistics)

There is another factor causing the variation in the variable under study. They are not regular variations and are purely random or irregular called residual variation. These variations are unforeseen, uncontrollable, and unpredictable. These forces are a rise in prices of steel due to strike in the factory, accident due to failure of the break, earthquakes, wars, flood, famines, and any other disasters. Irregular fluctuation is revealed in the figure 4.

Decomposition of Components

Decomposition models are of three types: additive, multiplicative, and pseudo-additive model.

(i) ***Additive Model:*** In some time series, the dimension of both the seasonal and irregular variations does not alter as the level of the trend rises or falls, where, an additive model is appropriate. In the additive model, the observed time series (O_t) is considered to be the sum of three independent components: the seasonal S_t, the trend T_t and the irregular I_t. Mathematically, it can be expressed as $O_t = S_t + T_t + I_t$

(ii) ***Multiplicative Model:*** The extent of both the seasonal and irregular variations increase as the level of the trend rises in many time series. In this situation, a multiplicative model is usually appropriate. In this model, the original time series is expressed as the product of trend, seasonal and irregular components. Mathematically, it can be expressed as $O_t = (S_t \times T_t \times I_t)$

(iii) ***Pseudo—additive Model:*** This model, normally, is not used when the original time series have very small or zero values as it is not possible to divide a number by zero. Against those cases, a pseudo-additive model combining the elements of both the additive and multiplicative models is used assuming that seasonal and irregular variations are both dependent on the level of the trend but independent of each other. The original data is expressed mathematically as $O_t = [T_t + T_t \times (S_t - 1) + T_t \times (I_t - 1)]$. $O_t = [T_t \times (S_t + I_t - 1)]$.

This model continues the convention of the multiplicative model to have both the seasonal factor S_t and the irregular factor I_t centred around one. Therefore, there is a need for subtraction one from S_t and I_t in order to ensure that the terms $[T_t \times (S_t - 1)]$ and $[T_t \times (I_t - 1)]$ are centred around zero.

2. Box-Jenkins Methodology: Econometric Forecasting

Forecasting is the process of making predictions of the future based on past and present data in analyzing trends. In this section, an attempt has been made to discuss one of the most popular approaches towards econometric forecasting which is known as Autoregressive Integrated Moving Average (ARIMA) forecasting method.

- Box and Jenkins introduced first the ARIMA model for short-term forecast of univariate or multivariate time series in 1976 and hence this method is known as the Box-Jenkins methodology. ARIMA is a combination of AR (*Autoregressive*), I (*Integrated*), and MA (*Moving Average*) process.
- A convenient notation for ARIMA model is ARIMA (p, d, q). Here, p, d, and q are the levels for each of the AR, I, and MA process. Thus, given the values of p, d, and q, it can be said that what process is being modelled.
- Each of these three processes is an effort to make the final residuals displaying a white noise pattern (or no pattern at all).
- Before conducting the ARIMA model, it is necessary to check that the underlying time series are stationary or they can be made stationary with appropriate transformations.
- Forecasting involves evaluating quantitative data for a forthcoming period based on past data incorporated in the model. This model is applied for policy-makers for forecasting economic variables, like GDP, population growth, employment, interest rate, inflation, job growth, production, consumption, etc. On the other hand, it helps companies and financial institutions in strategic planning.
- A time series can be modeled either in AR process, MA process or ARMA process.
- In AR process, current value of a variable depends upon the value of the variable in the previous period and an error term. So, if Y_t depends on its lagged value up to 'p' number of years (Y_{t-1}, Y_{t-2},…. Y_{t-p}), it is called AR (p) model.
- In order to estimate the model, it is necessary to evaluate autocorrelation function (ACF) (correlation of current value with its lagged values) and partial autocorrelation function (PACF) (correlation between an observation of 's' period ago and current observation after controlling for observations in intermediate lags) of the time series. In other words, the correlation between Y_t and Y_{t-k} after removing the effect of Y_{t-k+1} , Y_{t-k+2} ,……., Y_{t-1}
- In MA model, current value of a variable is determined based on current and immediate past values of the error term. MA (q) is simply a linear combination of white noises where Y_t depends upon current and previous values of white noise error term up to 'q' period (μt_{-1}, μ_{t-2}…μ_{t-q}).
- ACF and PACF of the MA model is estimated using AR process. If we plot MA (q) with the help of a Correlogram, the ACF and PACF will show a declining trend asymptotically reaching '0' with 'q' significant spikes.
- If we think that Y_t has characteristics of both AR and MA, it is called ARMA model where current value of Y_t in ARMA (p, q) depends upon lagged values of Y_t for 'p' period and μ_t for 'q' period. AR, MA and ARMA are applied if the series is stationary.
- However, if the series is integrated of order 'd', the original time series is said to follow ARIMA (p, d, q).
- After appropriate transformation of the series, the series is to be made stationary and ARMA (p, q) is to be applied on it *[Example: Series ARIMA (2, 1, 2); take first difference of the series; apply ARMA (2, 2)]*.
- ARIMA (p, 0, q) = ARMA (p, q); ARIMA (0, 0, q) = MA (q); ARIMA (p, 0, 0) = AR (p).

- Hence, it is important to identify whether a time series follows AR, MA, ARMA or ARIMA. In order to identify the model, it is necessary to evaluate the values of p, d, and q.
- If the series is stationary (d=0), transformation is not necessary. Stationarity of the series can be obtained by running unit root tests. After getting the value of 'd', number of times, it is to be differenced to make the series stationary can be obtained.
- In order to evaluate the value of 'p' and 'q', we need to formulate a Correlogram (plots of ACF and PACF with corresponding lags). For a MA (q) process, the ACF will show spikes (significant ACFs) up to 'q' lag whereas the PACF for MA (q) will tend to die down quickly either by an exponential decay or a damped sinewave.
- On the contrary, for AR (p) model, the PACF will show spikes (significant PACFs) up to 'p' lag whereas the ACF for AR (p) will tend to die down quickly either by an exponential decay or a damped sinewave.
- A mixed process is required when none of ACFs and PACFs show a definite cut-off. *[Example: if both ACF and PACF show exponential decay at lag 1, ARMA (1, 1) to be applied; if ACF show significant spikes up to lag 3, and PACF begins exponential decay after lag 1, ARMA (3,1) model may be applied.]*
- Using ordinary least squares (OLS) or maximum likelihood method, the original series Y_t is evaluated and forecasting is made.
- However, the BJ forecasting method consists of following steps:

Step 1: Identification of best fit ARIMA model
  (a) Plot time series data
  (b) Difference data to make stationary on mean (remove trend)
  (c) Log transform data to make data stationary on variance
  (d) Difference log transform data to make data stationary on both mean and variance
  (e) Plot ACF and PACF to identify potential AR and MA model

Step 2: Estimation of the best fitted model

Step 3: Diagnostics checking, i.e., finding out whether estimated residuals are white noise. If yes, go to the next step. If no, return to steps 1.

Step 4: Forecasting using the best fit ARIMA model

Step 5: Forecast Evaluation (goodness of fit)

> *Example: Consider a data series of wheat production for the period 1970-71 to 2010-11 in India. In order to forecast wheat production beyond 2010-11, using ARIMA method, we obtain stationarity of the series using ADF test. The series is found to be I (1). Hence, the series is differenced of the first order to make it stationary. Then the Correlogram of the series is plotted where ACF exhibits a significant spike at lag 1 but no clear pattern after that point. On the other hand, PACF shows significant spikes at lag 1 and 10. Hence, the series may be modeled at ARIMA (1, 10). Applying these values in econometric software like E-views, we will be able to forecast the series beyond 2010-11.*

3. Evaluation of Forecast Accuracy

After estimation of ARIMA model, it is necessary to evaluate the accuracy of the forecasting model and to choose between different ARMA models when different alternatives are available.

- In order for a model to be accurate, the residuals should be stationary (white noise); parameters should be statistically significant; and the model should be fit on the past data.
- Stationarity of the residuals can be tested with the help of ADF test. Statistical significance of the parameters are tested using 't' or 'z' test.

- Now, a model can be best fit of the past data by adding additional explanatory variables. However, the model will lack its forecasting ability. Hence, AIC or SBC may be used for finding out appropriate parameters of the model with robust forecasting ability.

Example: In the previous example of wheat production, an idea about reliability of the forecast can be obtained on the basis of following aspects: (a) the estimated coefficients of AR (1) and MA (10) are statistically significant at 1% level; (b) All inverted AR and MA roots are within unit circle (<1) which implies that the ARIMA structure is stationary and the model is correctly specified; (c) the value of root mean squared error (RMSE) is low; (d) values of bias proportion and variance proportion is low and value of covariance proportion is high (obtained from software). Hence, the model has good forecasting ability.

4. Non-Stationarity and Unit Root Test

There are important differences between stationary and non-stationary time series.

- In stationary time series, shocks will be temporary and over time their effects will be eliminated as the series revert to their long-run mean values. On the other hand, non-stationary time series will necessarily contain permanent components. Therefore, the mean and/or the variance of a non-stationary time series will depend on time, which leads to cases where a series (a) has no long-run mean to which the series returns, and (b) the variance will depend on time and will approach infinity as time goes to infinity.
- Stationarity of a time series denotes *means and variances a time series variable is constant over time and covariance between two time periods will depend on the distance between two periods and not the actual time.* A variable with these characteristics is all weakly or covariance stationary.
- However, if all the moments of the probability distribution (not just mean and variance) are constant over time, it will be called strictly stationary. If the data is not stationary, forecasting with such data is not practical.
- There are different mechanisms of testing stationarity of a series. Unit root test is one of them. A formal test of stationarity is unit root test.
- If a time series depends on its one year lagged series and noise term it can be represented as follows:

  $\Rightarrow$ $\mathbf{Y_t = \rho Y_{t-1} + u_t}$**, if** $\boldsymbol{\rho = 1}$

- If $\rho = 1$, the series is said to have unit root and is non-stationary.
- Thus, the null hypothesis to test the stationarity of the series is as follows:

  $\Rightarrow$ $H_0: \rho = 1$ ***or*** $\rho - 1 = 0$ ***or*** $\delta = 0$ ***where*** $\delta = \rho - 1$

  $\Rightarrow$ $H_1: \delta < 0$

- If it is not statistically significant, the series is said to have unit root and it is non-stationary. However, t-test cannot be applied to test the significance.

Unit Root Test

Hence, Dickey-Fuller (DF) or Augmented Dickey-Fuller (ADF) test (tau test), Phillips-Perron (PP) test are applied for unit root test.

(a) DF and ADF test

The null hypothesis to test the stationarity of a series is as follows:

$\Rightarrow$ $H_0: \rho = 1$ ***or*** $\rho - 1 = 0$ ***or*** $\delta = 0$ ***where*** $\delta = \rho - 1$

$\Rightarrow$ $H_1: \delta < 0$

In order to test the above hypothesis, Dicky-Fuller (DF) test is applied (Dickey & Fuller 1979). The test statistics (τ) for the above test is

$$\boldsymbol{\tau = (\hat{\delta} - \delta)/\,SE\,(\hat{\delta})}$$

Here, the assumption is the error term variance is constant and they are serially independent. The calculated value of τ is compared to its critical value using Monte Carlo experiments (McKinnon, 1991).

Decision Rule

- *If the absolute value of test statistic is greater than its critical value, H_0 is rejected and series is a stationary series and vice versa. ADF test is same as that of DF test with the exception that the test equation is augmented by including 'm' lags of dependent variable (ΔY_t) to correct any serial correlation in the disturbance term.*
- In case of ADF test, the model incorporate 'm' lags to avoid any serial correlation in the disturbance term. Calculation of 'm' may be done based on nature of data:
  - ♦ *m=1 or 2 in case of annual data*
  - ♦ *m=12 in case of monthly data*
  - ♦ *m=integer part of $12\times(T\div100)^{1/4}$*
  - ♦ *ADF is conducted by setting $m=m_{max}$. The particular lag where the absolute value of t-statistic is more than 1.6 is the level of 'm'.*

DF test or ADF test can be performed in three different forms:

- *Random walk (None):* If a time series fluctuates around a sample average of zero, it is called random walk model without a drift (i.e. no intercept). Mathematically it may be represented as follows: $\boldsymbol{\Delta Y_t = \delta\, Y_{t-1} + u_t}$ in case of DF test and $\boldsymbol{\Delta Y_t = \delta\, Y_{t-1} + \sum_{i=1}^{m} \gamma_i\, \Delta Y_{t-i} + u_t}$ in case of ADF test.
- *Random walk with drift (constant/intercept):* If the time series fluctuates around a sample average that is nonzero, it is called random walk with drift (i.e. with an intercept). Mathematically it may be written as: $\boldsymbol{\Delta Y_t = \alpha + \delta\, Y_{t-1} + u_t}$ where $\boldsymbol{\alpha}$ being the drift or shift parameter in case of DF test and $\boldsymbol{\Delta Y_t = \alpha + \delta\, Y_{t-1} + \sum_{i=1}^{m} \gamma_i\, \Delta Y_{t-i} + u_t}$ in case of ADF test.
- *Random walk with drift around a deterministic trend (constant and trend):* If the time series fluctuates around a linear or quadratic trend, it is called random walk with drift and deterministic trend. Mathematically it may be represented as $\boldsymbol{\Delta Y_t = \alpha + \beta t + \delta\, Y_{t-1} + u_t}$ in case of DF test and $\boldsymbol{\Delta Y_t = \alpha + \beta t + \delta\, Y_{t-1} + \sum_{i=1}^{m} \gamma_i\, \Delta Y_{t-i} + u_t}$ in case of ADF test**.** Deterministic trend value, β is added with each time period

With a view to selecting the appropriate model for the test, the following methodology may be followed:

(i) For this purpose, the data series is to be plotted graphically. If no visual trend is observed in the data series, the data series may be RW with a drift (intercept).
  - With a view to estimating this, ADF test is conducted with 'none' model first. If H_0 ***(ρ = 1)*** is accepted, it is concluded that the series is a pure RW (non-stationary) series that follows only stochastic trend.
  - However, if it is rejected, ADF test is conducted again with 'intercept' model. If H_0 is accepted now, it is concluded that the series is RW with a drift. However, if it is rejected then the series is a stationary series.

(ii) On the other hand, if a visual trend is observed in the series, the series may be RW with trend and drift.

- With a view to estimating the same, ADF test is conducted with 'intercept' model first. If the H_0 is accepted, the pre-described procedure is to be followed to estimate whether a series is RW with a drift.
- However, if H_0 is rejected, ADF test is to be conducted once again with 'trend and intercept' model. If H_0 is accepted now, the series is RW with trend and drift (stochastic as well as deterministic trend). But if it is rejected, the series is a stationary series.

Example: If we consider a daily time series data of SENSEX from 1992 to 2018, stationarity of the series using ADF and PP test can be estimated as follows:

Results of ADF and PP Test

Test Series	Model	Test statistics	Critical values 1%	5%	10%	Prob.	Decision rule based on critical values	Decision rule based on Prob.	Decision on H_0	Remarks
ADF test (at level)										
LSENSEX	Intercept	-0.134051	-3.431272	-2.861832	-2.566968	0.9440	Absolute value of test statistics<Absolute critical values at 1%, 5% and 10%	Prob.>0.10>0.05 >0.01	Accepted	The series has unit root
LSENSEX	Trend and Intercept	-3.05713	-3.95958	-3.41056	-3.12705	0.1169	Absolute value of test statistics<Absolute critical values at 1%, 5% and 10%	Prob.>0.10>0.05 >0.01	Accepted	The series has unit root
LSENSEX	None	1.493659	-2.56535	-1.94088	-1.61666	0.9671	Absolute value of test statistics<Absolute critical values at 1%, 5% and 10%	Prob.>0.10>0.05 >0.01	Accepted	The series has unit root
ADF Test at First Difference										
DLSENSEX	Intercept	-72.2447	-3.43127	-2.86183	-2.56697	0.0001	Absolute value of test statistics>Absolute critical values at 1%, 5% and 10%	0.10>0.05>0.01> Prob.	Rejected	The series does not have unit root
DLSENSEX	Trend and Intercept	-72.2504	-3.95958	-3.41056	-3.12705	0.0000	Absolute value of test statistics>Absolute critical values at 1%, 5% and 10%	0.10>0.05>0.01> Prob.	Rejected	The series does not have unit root
DLSENSEX	None	-72.2211	-2.56535	-1.94088	-1.61666	0.0001	Absolute value of test statistics>Absolute critical values at 1%	0.10>0.05>0.01> Prob.	Rejected	The series does not have unit root
PP Test at Level										
LSENSEX	Intercept	-0.0994	-3.4313	-2.8618	-2.567	0.9478	Absolute value of test statistics<Absolute critical values at 1%, 5% and 10%	Prob.>0.10>0.05 >0.01	Accepted	The series has unit root
LSENSEX	Trend and Intercept	-3.1182	-3.9596	-3.4106	-3.1271	0.1020	Absolute value of test statistics<Absolute critical values at 1%, 5% and 10%	Prob.>0.10>0.05 >0.01	Accepted	The series has unit root
LSENSEX	None	1.484391	-2.5654	-1.9409	-1.6167	0.9664	Absolute value of test statistics<Absolute critical values at 1%, 5% and 10%	Prob.>0.10>0.05 >0.01	Accepted	The series has unit root
PP test at First Difference										
DLSENSEX	Intercept	-72.177	-3.43127	-2.86183	-2.56697	0.0001	Absolute value of test statistics>Absolute critical values at 1%, 5% and 10%	0.10>0.05>0.01> Prob.	Rejected	The series does not have unit root
DLSENSEX	Trend and Intercept	-72.1805	-3.95958	-3.41056	-3.12705	0.0000	Absolute value of test statistics>Absolute critical values at 1%, 5% and 10%	0.10>0.05>0.01> Prob.	Rejected	The series does not have unit root
DLSENSEX	None	-72.1672	-2.56535	-1.94088	-1.61666	0.0001	Absolute value of test statistics>Absolute critical values at 1%	0.10>0.05>0.01> Prob.	Rejected	The series does not have unit root

(b) Phillip-Perron (PP) test

- This model is actually a modification over ADF model. Here, the serial correlation in the error term is corrected while calculating the test statistic. The procedure for performing PP test is same as that of ADF test.
- However, while ADF test is based on the assumption of serial independence of the error term, PP test is less restrictive about this assumption. The test statistics in the test is corrected accordingly. The asymptotic distribution of the test statistic in PP test is same as that of ADF test (Phillips & Perron, 1988).
- If a series is not stationary at level, a first difference of the series is to be calculated by taking 1-year lag of the series.

5. Granger Causality

The test basically measures the direction of relationship between two variables. It is a statistical test that is based on the premise that future cannot cause the present or past. In an analysis involving two or more variables, if one variable occurs before the others, the variable is said to cause the other variables. So, the Granger causality test identifies whether one variable precedes the other or they are contemporaneous.

- If only variable granger causes the other, it is called unidirectional causality. However, if both the variables granger causes one another, it is called bidirectional causality.
- If none of the variable granger causes the other, they are said to be independent of one another.
- If Variable-A granger causes Variable -B and Variable -B granger causes Variable -C, it is not necessary that Variable -A granger causes Variable -C. It may create a spurious causality.
- With a view to conducting Granger causality test, one variable is regressed against the lagged values of the same variable without incorporating the other variable in the equation.
- Then residual sum of square (RSS) of this regression is estimated which is termed as RSS_R. Once again the same variable is regressed against the lagged values of that variable and the other variable under consideration. The RSS of this regression is labelled as RSS_{UR}.
- Causality among the variable is tested based on the following hypothesis:
  - ⇒ ***H_0: One variable does not granger cause the other variable***
  - ⇒ ***H_1: One variable granger causes the other variable***

The H_0 is tested based on following test statistics: $F = (RSS_R - RSS_{UR}) \times (n-k)/RSS_{UR} \times m$, *Where, m = number of lagged terms; k = number of parameters estimated in the unrestricted model.*

Decision Rule

- At (m, n—k) degrees of freedom and 5% level of significance, if the probability of obtaining calculated of F in the F distribution table is less than 0.05, H_0 is rejected and vice versa.
- If H_0 is rejected, the one variable granger causes the other. The variables for Granger causality test have to be stationary.
- If they are not, the observations under the variables are differenced to make them stationary.

Selection of lags is important in deciding the causality between two or more variables. Hence, lags are to be decided based on AIC or SBC criterion discussed below. If there are three variables in the study and the first variable granger causes the second and the second granger causes the third, it does not necessarily mean that the first variable would granger cause the third. Any relation between the first and third variable might be spurious in nature.

Example: There are 3 variables – Sensex, S&P 500 and SSE Composite Index. Based on daily data of these three indices, their granger causality is to be estimated. However, for the purpose, the data must be stationary. The log-transformed data was non-stationary at level. Hence, they were made stationary at first difference.

Null Hypothesis (H_0)	F-Statistic	Prob.	Decision rule	Decision on H_0	Remarks
D(LSP) does not Granger Cause D(LSENSEX)	45.2945	2.E-37	Prob.<0.05	Rejected	D(LSP) Granger Causes D(LSENSEX)
D(LSENSEX) does not Granger Cause D(LSP)	0.48339	0.7480	Prob.>0.05	Accepted	D(LSENSEX) does not Granger Cause D(LSP)
D(LSSE) does not Granger Cause D(LSENSEX)	1.00408	0.4039	Prob.>0.05	Accepted	D(LSSE) does not Granger Cause D(LSENSEX)
D(LSENSEX) does not Granger Cause D(LSSE)	4.40543	0.0015	Prob.<0.05	Rejected	D(LSENSEX) Granger causes D(LSSE)
D(LSSE) does not Granger Cause D(LSP)	1.26709	0.2805	Prob.>0.05	Accepted	D(LSSE) does not Granger Cause D(LSP)
D(LSP) does not Granger Cause D(LSSE)	8.41210	9.E-07	Prob.<0.05	Rejected	D(LSP) Granger Causes D(LSSE)

There is unidirectional causality between each pair of indices. From the result, it may be finally inferred that the movement of S&P 500 influences the movement of both SENSEX and SSE Composite Index but movement of SENSEX only influences the movement of SSE Composite Index. However, SSE Composite Index cannot influence any one of the two indices.

(a) Selection of Appropriate Lag

The methods used for selection of appropriate lags in a VAR model are – (a) Akaike Information Criterion (AIC); (b) Schwarz Bayesian Criterion (SBC); and (c) Hannan-Quinn Criterion (HQC). In the context of linear models, the criterions are computed as follows:

- ⇒ $AIC = e^{2k/n} (RSS/n)$
- ⇒ $SBC = e^{k/n} (RSS/n)$
- ⇒ $HQC = (\ln n)^{2k/n} (RSS/n)$

Where, RSS = residual sum of squares; k = number of parameters estimated (including the intercept); and n= number of observations.

- While selecting an appropriate lag for the VAR model, the lag that reports the minimum value in these criterions is to be selected as the appropriate one.
- However, if there is any contradiction between AIC and SBC model with respect to selection of appropriate lag, the lag selected as per HQC model should be the most appropriate one.

Lag	AIC	SC	HQ
0	3.269223	3.272606	3.270399
1	-16.03079	-16.01726	-16.02609
2	-16.06797	-16.04429*	-16.05974
3	-16.07179	-16.03796	-16.06004
4	-16.07753	-16.03356	**-16.06225***
5	-16.07887	-16.02475	-16.06006
6	-16.07980	-16.01552	-16.05747
7	-16.07981*	-16.00539	-16.05395
8	-16.07821	-15.99363	-16.04882

In the above example, Min (AIC) and Min (SBC) occurred at Lag 7 and Lag 2 respectively. Hence, the lag at Min (HQC) (Lag 4) is to be selected as the appropriate lag.

(b) Granger causality in Vector Autoregressive (VAR)

- In VAR model, there is no prior distinction between endogenous and exogenous variable (Sims, 1972). Hence, VAR involves simultaneous equations where value of a variable is represented as a linear function of the lagged values of that variable and all other variable included in the model.
- With a view to estimating VAR model, the variables must be stationary. If they are not stationary, they should be made stationary by differencing, and the differenced variables are to be used in the model.
- Based on appropriate lag length, the equations in the VAR model is estimated by applying ordinary least square (OLS) technique. The coefficient (β_T) of each equation in the model is estimated as follows:

  ⇒ $\beta_T = Cov(X_T, Y_T) / Var(X_T)$

- While the magnitude and sign of the coefficient determines the degree and nature of relationships of a variable with its own lagged values and lagged values of other variables, significance of such relationships is estimated using t-test based on following hypothesis:

  ⇒ $H_0 : \beta_T = 0$

  ⇒ $H_1 : \beta_T \neq 0$

The test statistics for the test can be formulated as follows:

⇒ ***t = un-standardised β_T / standard error (SE) of β_T***

- At (n-2) degrees of freedom and 5% level of significance, if the calculated absolute value of t is less than its tabulated value, H_0 is accepted and vice versa.
- If H_0 is accepted, the particular independent variable in that equation does not have significant relationship with the dependent variable of the same equation.

VAR model allows testing the direction of causality between variables. Hence, Granger causality test can be performed using the VAR equations. It is based on testing the joint significance of the lags of each variable, apart from its own lags. Hence, in this test, causality among two variables is analysed based on following hypothesis:

⇒ ***H_0 : one variable granger causes the other***

⇒ ***H_1 : one variable does not granger cause the other***

Decision Rule

- The test statistic follows Chi-square distribution.
- At 5% level of significance, if the probability of obtaining the test statistic in Chi-square distribution table is less 0.05, H_0 is rejected and vice versa.
- If H_0 is rejected, one variable does not granger cause the other. However, the reverse may not be true.
- If there are more than two variables in a VAR model, Granger causality in VAR is conducted to estimate causality between each pair of variables individually and causality of all the variables jointly with one variable.

> *Example: There are 3 variables – Sensex, S&P 500 and SSE. Current value of Sensex is represented with the help of lagged values of Sensex, S&P 500 and SSE. Now, if we want to see whether S&P 500 granger causes Sensex, we need to estimate the significance of the coefficients of lagged values of S&P 500 in the first model. If they are significant, then S&P 500 granger causes Sensex. In that way, pair-wise granger causality can be estimated for each pairs.*

Dependent Variable	Excluded	Chi-sq	df	Prob.	Decision rule	Decision
D(LSENSEX)	D(LSP)	182.1615	4	0.0000	Prob.<0.05	D(LSP) Granger causes D(LSENSEX)
	D(LSSE)	5.087174	4	0.2785	Prob.>0.05	D(LSSE) does not Granger cause D(LSENSEX)
	All	186.2986	8	0.0000	Prob.<0.05	D(LSP) and D(LSSE) jointly Granger cause D(LSENSEX)
D(LSP)	D(LSENSEX)	4.024418	4	0.4027	Prob.<0.05	D(LSENSEX) does not Granger cause D(LSP)
	D(LSSE)	7.158203	4	0.1278	Prob.>0.05	D(LSSE) does not Granger cause D(LSP)
	All	9.092785	8	0.3345	Prob.>0.05	D(LSENSEX) and D(LSSE) jointly do not Granger cause D(LSP)
D(LSSE)	D(LSENSEX)	10.13211	4	0.0383	Prob.<0.05	D(LSENSEX) Granger causes D(LSSE)
	D(LSP)	26.12782	4	0.0000	Prob.<0.05	D(LSP)Granger causes D(LSSE)
	All	43.81535	8	0.0000	Prob.<0.05	D(LSENSEX) and D(LSP) jointly Granger cause D(LSSE)

S&P 500 Granger causes SENSEX, while SSE Composite Index does not do so. However, both these indices together Granger causes SENSEX. On the other hand, neither SENSEX, nor SSE Composite Index nor these two indices together separately Granger causes S&P 500. SENSEX, S&P 500 and these two indices together separately Granger causes SSE Composite Index.

(c) Granger causality in VECM

If the variables are not stationary but co-integrated, vector error correction model (VECM) can be applied to estimate the relationship among variables. Causality among them in this situation can be estimated using granger causality in VECM using the same procedure as that of granger causality in VAR.

Note: In case of Granger causality in VAR or VECM, one series is considered as endogenous, while other series or the other 2 series are considered as exogenous. That's why these tests are also known as Box Exogeneity test. That's the reason why causality of 2 series on 1 series can be computed in this method. However, in case of Granger causality, all the data series are considered to be endogenous. Hence, their relationships are always 1-1

- The similarity between these 2 tests is that the data must be stationary in both the cases. It is different from that of granger causality in VECM, where the requirement is non-stationarity.

- *While the underlying test for granger causality is F-test, it is Chi-square in case of VAR granger causality.*
- The basic difference between these 2 tests is that in case of granger causality, only 1 equation is formulated to explain causality among 2 or more data series.
- However, in case of VAR granger causality, the number of equation = number of data series.
- Suppose there are 3 series in a study. Then the equations to be formulated are: Series A = f(Series B); Series A = f(Series C); and Series A = f(Series B and Series C).
- As one series is taken as exogenous variable, hence VAR granger causality is also known as block Exogeneity test.

6. Co-integration and Error Correction Mechanism

- If the time series variables are stationary, their relationships can be estimated using OLS technique.
- But if they are non-stationary, the technique to be applied is called cointegration.
- In fact, if a data series is non-stationary, they can be made stationary by taking the first differences. But it may lead to potential data loss.
- The short term and long term dimensions of relationships will also not be identified in OLS technique.
- Hence, we consider two time series data which are non-stationary and integrated of order d, where d = 0, 1, 2, etc. If d=0, the data are already stationary. If d=1, the data can be made stationary by taking first difference. If d=2, the data can be made stationary by taking the second difference and so on.
- Normally, in case of co-integration, the time series are non-stationary with integrated order 1[I(1)]. However, in case of inflation, money supply, it may be I (2). If two series are non-stationary with I (1), their linear combination will be stationary [I (0)].
- Examples where two series may have cointegration relationships:
  - *Short and long term interest rate*
  - *Spot and future prices of products*
  - *Prices and wages*
  - *Price at home and abroad*
- If two or more series are not integrated of same order, cointegration is not possible. In that case, *Autoregressive Distributed Lag (ARDL)* method is to be applied.

Engel Granger (EG) Co-integration Test

- In order to check cointegration between two series, Engel Granger (EG) Cointegration test is conducted. As per this test, stationarity of two series are tested and if both the series are found to be I (1), a linear combination of the two series is made using OLS technique and residuals of the model are estimated.
- Stationarity of the estimated residuals is tested. If the estimated residuals are found to be stationary, the series is said to be cointegrated. This is called Engle-Granger cointegration. There are some problems to the EG approach:
  - *It cannot be used in models with more than 2 variables as it may lead to creation of more than 1 cointegrating vectors.*
  - *As estimated residuals are calculated based on OLS, there is a possibility of further error as the series is non-stationary.*
  - *Change in the regressor (independent variable) in the linear combination of two non-stationary series may change the result.*
  - *If there is structural break in the series, this method cannot be applied.*

Johansen Co-integration Test

- In order to do away with those problems to the EG approach, Johansen Cointegration test is applied. However, *cointegration only estimates the long term relationships among variables.*
- In order to capture *both short term and long term relationships*, Error Correction Model (ECM) is applied. In this model, first difference in one of the variable is expressed in terms of first difference in the other variable, one year lagged information of estimated residual and error term.
- If one year lag of estimated residual is non-zero (positive or negative), there is disequilibrium in the short run. Relationship between two cointegrating variables is estimated based on ECM model.
- If the residual in the model is negative and significant, short run disturbances can be modified by increasing or decreasing the variables in the current period to ensure long run equilibrium.
- As the data in ECM model are stationary, the coefficients can be estimated using OLS and their significances may be tested using t-test. However, when the variables are non-stationary and cointegrated, VECM is to be applied.

Example: There are 3 variables – Sensex, S&P 500 and SSE Composite Index. They are all I (1) series. Cointegration among the variables can be estimated using JJ cointegration test and their internal relationship can be estimated using VECM model. Under Trace statistics and Max Eigen Value Statistics, the result of cointegration is as follows:

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	**0.004583**	**36.76827**	**35.19275**	**0.0335**
At most 1	0.001008	9.509994	20.26184	0.6877
At most 2	0.000594	3.525140	9.164546	0.4874

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	**0.004583**	**27.25828**	**22.29962**	**0.0094**
At most 1	0.001008	5.984853	15.89210	0.7896
At most 2	0.000594	3.525140	9.164546	0.4874

As hypothesized number of cointegrating equation none is rejected, all three series are cointegrated.

Note 1 (Cointegration and VECM):

- Co-integration means change in one series is a function of change in another series. So, if the series is co-integrated, it means that two or more series would converge in the long run.
- In fact, they have long run equilibrium. However, in order to establish their relationship, we need to frame a model. The error in the model is required to be corrected continuously to achieve the equilibrium position.
- That's why the model is called Vector Error Correction Model. This equilibrium is technically called short term equilibrium.
- The causality among the series is based on this VECM model. That's why they are also called short term causality.
- *There is no such thing as short run cointegration. Cointegration is always meant to find out the long term behaviour of the data series.*
- Now, long term does not mean any particular time period. If we conduct cointegration based on daily data of 15 years, it will also project long term behaviour.
- If we do it with daily data of only 5 years, that will also project long term behaviour.

- In fact, based on a sample of 15 years or 5 years, we may conclude whether the data series influence one another over a long period of time (mainly appropriate for taking monetary decisions of the Government).
- But what is the nature of such influence is not known from cointegration. That's why VECM is conducted so that it restricts long term behaviour of the endogenous variables while allowing for short run adjustment dynamics.
- *In case of non-stationary data, short run relationships (important for portfolio diversification) are estimated through granger causality in VECM as discussed in previous segment.*

Note 2 (Structure of Analysis):

- First, check for the stationarity in the data series.
- If all the series are stationary, go for VAR for estimating the relationship among them.
- If the series are not stationary, check whether the series are integrated of same order. If they are, go for cointegration.
- If cointegration exists among the series, their relationship is estimated using VECM.
- If there is no co-integration, the series are to be converted into a stationary series taking first difference and then VAR is to be applied on the first differenced data.
- If the series are not integrated of the same order, apply ARDL.

7. Volatility Modeling

- Financial time series are subject to volatility clustering. It means variability in the values for an extended period followed by a relative calm. This volatility clustering can be seen in inflation data (may create problem for policy decision making); exchange rates (may create problem for exporters and importers); stock prices (may create problems for investors); capital market (may create problems for issuers in raising money).
- Usually financial time series follows random walk at level form and hence, non-stationary which can be converted into stationary by taking first difference. However, this first-differenced data may also have unequal variances over time. It is called **Autoregressive Conditional Heteroscedasticity (ARCH)** effect (Engle, 1982). If the unequal variances (Heteroscedasticity) of the time series are autocorrelated (autoregressive), it is said to have ARCH effect and hence subject to volatility clustering in data.
- If a time series Y_t is a linear function of 'k' variables, we assume that the disturbance u_t has a mean '0' and Var (u_t) is a function of lagged squared disturbances up to period 'p' [u^2_{t-p}] (also known as variance equation). In order to test ARCH effect, we take the following hypothesis:
  - ⇒ **H_0: lagged squared disturbances are not autocorrelated**
  - ⇒ **H_1: lagged squared disturbances are not autocorrelated**
- F-test is conducted to test the above hypothesis. In order to test the above hypothesis, the test statistics $T \times R^2$ is calculated, where T is the number of observations, and R^2 is the coefficient of determination of the variance equation in the ancillary regression equation replacing u with the (Actual Y_t —Estimated Y_t).
- For large samples, test statistics follows chi-square distribution. If calculated value of test statistics is > its tabulated value at chi-square distribution table at chosen level of significance and 'p' degrees of freedom, H_0 is rejected and vice versa. ARCH model may also be used in AR or ARDL model.

Example: Consider a daily % change in SENSEX as the financial time series. The graphical representation of the data shows volatility clustering at some period with periods of relative calm subsequently. We estimate the AR (1) model of the data (current value depends upon 1 year lagged value). Then we run the ARCH test on the AR (1) model with lag 1. Since p-value<0.01, H_0 is rejected and ARCH effect is present.

8. GARCH Model

Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model developed by Bollerslev (1986) is an improvement over ARCH model. In a GARCH (p, q) model, variance of the disturbance at time 't' (u_t) is represented as a function of lagged squared disturbance up to previous period 'p' (u^2_{t-p}) and lagged conditional variance up to previous period 'q' (σ^2_{t-q}).

- If we set p=1 and q=0, GARCH (p, q) gets reduced to ARCH (1). Hence, if all the coefficients of the conditional variance are 0, GARCH (p, q) gets reduced to ARCH (p) model.
- GARCH model is very flexible and covers wide variety of financial volatility. The model is estimated by maximum likelihood method.
- In case of large samples, it follows normal distribution represented by t-statistics with confidence interval ± S.E.

Example: In the aforesaid example, GARCH (1, 1) is applied and coefficient of both lagged squared disturbance and lagged conditional variance are found to be statistically significant indicating volatility clustering. At the same time, sum of the coefficients are close to unity which means that the volatility shocks are quite persistent.

****