CHAPTER 1
INTRODUCTION

1.1. The Background of Decision Theory

Making decisions is an inevitable activity in life, whether at a personal level or at an institutional level. Everyone is faced with situations where a decision has to be made. There are two ways of treating such situations. One way is to consider the situation to be posing a challenge, where one is more worried about consequences of making a wrong decision. The other, obviously, is to consider the situation to be offering an opportunity, where one is interested in maximizing the benefits by making the right decision. Decision can be made for different purposes: from a simple personal decision like whether to carry an umbrella when going for a walk to a national policy decision like whether to abolish the death penalty when dealing with criminal justice. Decisions may be made at two levels - individual level and organizational. Individual decision may affect only the individual making the decision or may affect other people, depending on the nature of the decision. Similarly, organizational decisions may affect only the concerned organization or may affect many more individuals and organizations. The latter cases make decision making an important activity, both for individuals and organizations. Humans face decision making situations throughout their lives. It is therefore required to comprehend the very process of decision making on the part of individuals, in particular. There is a vast literature on the history of decision making, from the time it was considered to be a privilege of rulers to the time it was considered to be a right of citizens in a democratic setup. There are discussions about whether decision making is an art or a science. It may be interesting to know that there has also been a discussion on comparisons between choosing and deciding. Is making a choice same as making a decision? According to me there is a basic difference between making a choice and making a decision. In a recent book, Sheena Iyengar (2010) defines choice as "the ability to exercise control over ourselves and our environment." She further states: "In order to choose, it is first necessary to perceive that control is possible." Some of her conclusions are listed below:

- Choice is possible and desirable up to a certain point. A decision-maker gets confused beyond that point.
- Our health can be affected by choice on the job, partly depending on our need for choice.
It is very often found that any choice by the individual gets conditioned by various types of alternatives encountered by a decision-maker. It also gets affected by the people presenting these alternatives. This often happens when even when one alternative is clearly superior to others.

There has been a natural tendency to make irrational choices in an attempt to avoid loss as a result of making incorrect choice. It is likely to lead towards minimizing the loss of the potential.

Sometimes choices are made only to allow us to conform to the behavior or perception of others in relation to our perception of ourselves.

The individual choice is affected by the order in which he or she encounters options or alternatives. For example, in an interview given for a job, the first and the last person to be interviewed get benefitted. That is why traditional techniques of interviews are treated as one of the least useful tools for judging the future success of an employee.

The significance of any choice is culture-specific. Some cultures consider it more important than the other. It also differs from society to society. Choice in individualist societies is bound to be different from choice in collectivist societies. This means that no single approach is complete in itself in organizing and encouraging people.

Making choice can be difficult when the alternatives are almost tricky and equally good or bad distracting the individual making a choice.

Some researchers have concluded that making choice is often described as an art in order to justify a decision that has gone wrong. A successful decision-maker, on the other hand, calls his ability to make the right decision an art in order to emphasize that there is no system for making right decisions. Rather, he often claims that he makes his decisions on the basis of his gut feeling rather than on the basis of any scientific methodology. This very ambiguity of the process of decision making has made it the center of both curiosity and confusion. On one hand, everyone wants to know how to make right decisions. At the same, however, nobody wants to understand how decisions can be made more rationally, rather than based on intuition. The reason for this is that a good decision may not always be the best, but it will certainly be better than other decisions in the long run. Therefore, it is necessary to differentiate between situations that may occur frequently and conditions that may prevail very rarely. In the former case, it is reasonable to make a decision that
proves to be right on more occasions than it is wrong. In the latter case, there is no way to
develop a rule for making the right decision, because there may not be alternatives to
compare. Unprecedented climatic conditions or an earthquake are examples of the latter
case, while decisions like whether to renew an ongoing contract or whether to invest in a
certain stock are examples of the former case. Statistical science, being a science that
allows one to learn from experience, can address situations of the former nature, that is,
situations that occur frequently.

Historically speaking, the term ‘decision making’ was first used by Chester Barnard,
author of the book "The Functions of the Executives." It then became more common to use this
term in place of terms like "resource allocation" and "policy making" that were used more
commonly. William Starbuck, professor at University of Oregon, has argued that this phrase
changed the way managers thought about their functioning. He states that policy making could
go on and on endlessly, and there are resources to allocate. Decision implies the end of
deliberation and the beginning of action. This changed the perspective of managers from
deliberation to action. Albert Camus, a great philosopher, has stated: “Life is the sum of all your
choices.” It has been universally accepted that a scientific study of decision making becomes a
combination of several intellectual disciplines such as mathematics, statistics, logic, economics,
political science, psychology, sociology and many more. What do philosophers and historians do?
They just discuss and comment upon our decisions revealing ourselves and our values and
morals whereas historians list and objectively analyze the choices made by the leaders at critical
times. However, an organizational behaviour, originates from a more practical purpose. It
enables us to achieve better results. It is rightly believed that a good decision does not guarantee a
good result. Pragmatic approach in making decisions on the basis of merits, therefore, is likely to
pay off in the long run. The potential in managing risk in various contexts, a sensible perception
of human behaviour and the most advanced technology support and imitate innate cognitive
processes in human organisms, lead to the improved decision making capacity in many
challenging situations.

In spite of a significant progress in the process of decision making, we need to accept that
the history of decision making is not as smooth as it appears to be. Researchers have been well
aware of the limitations and constraints on the human capacity to make optimal right decisions.
Intricate circumstances, insufficient time, and most importantly, inadequate computational power
lead the decision makers to the state of bounded rationality. Though it has been argued that
people make rational and appropriate decisions if they collect enough information required, there
are various factors compelling the decision makers to decide against their material gains. In the
face of the inability to make perfect decisions, theoreticians have attempted to achieve at least
acceptable decisions.

1.1.1 Short Time Line of Decision Making

It is necessary to focus on the history of the process of decision making. A short time line of
decision making partially represents significant milestones in the history of decision making.
However, the study of the history of decision making is not complete in itself. It is only an
attempt to highlight important developments in the history of decision making. Developments in
the history of decision making are as follows:

- During prehistoric period for centuries, human decisions have been guided by so called
  supernatural and magical powers of the prophets and seers who were believed to have the
  power to see into the future.
- Lao-tzu taught and advocated the principle of non-willful action in the sixth century B.C.
  As per his recommendations, we should let events take their own natural course. Besides,
  Confucius recommended that decision making needs to be guided by nobility, benevolence,
  ritual, reciprocity, serenity and piety.
- Male citizens of Athens used to make decisions by voting in the fifth century B.C. It was
  the beginning of a democratic self-government in the true sense of the term.
- Plato, a great philosopher, strongly believed and asserted that all perceivable things are
  discovered in a better way only through the spiritual faculty rather than through the
  sensory faculty in the fourth century B.C. Aristotle, his disciple, did not agree with
  Plato and took an empirical view of knowledge that values information gathered only
  through human senses and deductive reasoning.
- Julius Caesar gave rise to the metaphor for making irreversible decisions in 49 B.C.
- The Hindu-Arabic number system, which included zero in it, introduced and initiated
  growth in the field of Mathematics in the ninth century.
- Omar Khayyam followed the Hindu-Arabic number system and created a language of
  calculations paving the way for growth and development of Algebra in the eleventh
  century.
An English friar proposed a rule of thumb for scientists in the fourteenth century. He maintained that the best theory is the simplest one that accounts for all the evidences. This theory has come to be popularly known as ‘Occam’s razor.’

Rene Descartes claimed that the reasoning power is superior to any experience for acquiring knowledge. He also established the framework for the scientific method in 1641.

Blaise Pascal and Pierre de Fermat jointly developed the concept of calculating probabilities of chance events in 1654.

In 1660, Pascal’s wager proved that the consequences of being wrong can be more important than the likelihood of being wrong for any decision maker. His has become a classic example of making a choice under uncertain circumstances.

In 1738, Daniel Bernoulli laid the foundation of risk analysis by assessing random events in order to examine the role individual desires or fears play in decision making. Thus the credit of laying the foundation of risk analysis by examining random events goes to him.

Carl Friedrich Gauss studied the bell shape curve in the nineteenth century. Previously, it was described by Abraham de Moivre who had also developed the structure for understanding the occurrences of random events.

In 1886, Francis Galton claimed that though the values of a random process may vary from the average, they are likely to tend towards it in the course of time. His concept of regression to the mean has influenced stock and business market analysis.

Frank Knight distinguished between risk and uncertainty in 1921. In risk the probability of an outcome can be known whereas in probability, the outcome is not known.

According to Abraham Wald there are two main procedures of sampling distribution based statistical theories. These theories are hypothesis testing and parameter estimation. These are special cases of the general decision problem. The paper he published in 1939, primarily deals with numerous concepts of statistical theory comprising loss functions, risk functions, admissible decision rules, antecedent distributions, Bayesian procedures, and minimax procedures. The term ‘decision theory’ was first used by E. L. Lehmann in 1950.

John von Neumann and Oskar Morgenstern wrote a book on the game theory in 1944 in which he has described a mathematical basis for economic and precise process of
decision making. They are of the view that decision makers, by and large, are rational and consistent.

- Howard Raiffa’s book, ‘Decision Analysis’, published in 1968, describes with illustrations decision-making techniques with special focus on the concept of decision trees and the expected value of the sample information.
- 1989. Howard Dresner has introduced and defined the term ‘Business Intelligence’ with the purpose of describing a set of methods leading to analytical decision making for improving business performance.

However, a question remains. The question is about the difference between choice and decision. Formally, the two terms are defined as follows:

A choice is a selection from a number or variety of options. It deals with possibilities and may lead to a position. In making a choice, one is concerned with the reward for the choice one makes. A decision is reaching a conclusion or passing judgment on an issue. It generates a direction and may lead to an action. In making a decision, one is concerned with the loss one may incur due to the decision made.

The present research is of statistical nature and goes more in the direction of decision making with a view to minimize the expected or potential loss as may be defined in the particular case that one is confronted with. Statistical decision theory is more quantitative in nature and uses mathematical tools to express the problem, the method, and the solution. Following is a brief review of the literature on statistical decision theory. It is not exhaustive because the literature on statistical theory cannot be reviewed in one chapter of a thesis. What is reviewed is the relevant literature from the point of view of the researcher, namely, data-driven decision making. Decision making is an integral part of human life, and more so of a professional life. Every time there is a dilemma, a decision is to be made. It is therefore natural that the decision theory is also as old as the need for resolving every situation involving a dilemma. This work considers the process of decision making as a tool to resolve a problem in complex situations.

The proposed thesis begins with a general theory of decision making. It then looks into the statistical decision theory. Since decision making is a human activity, it cannot be objective all the time. Hence it is therefore important to understand the objective and subjective sides of decision making. Another important consideration pertains to the complexity of the problem and hence of the decision. Simple situations may be resolved by simple decisions, but complex
situations may involve multiple decisions that may not be required or possible at one go. It is then a common practice to break up the problem in a sequence of simple questions that can be answered under prevailing conditions. This consideration leads to development of sequential decision rules. Since every stage of a sequential decision situation leads to several options, the situation can be represented using a tree diagram. It was in this way that the method of drawing decision trees was developed. Decision trees are easier to visualize than decision tables when the problem is complex. A tree representation also makes it possible to detect and identify exceptional cases, so that validity of the decision rule can be questioned or tested. This consideration has led the recent development of data-driven decision rules rather than rule-based decision rules. Another contrast is between data-driven decisions as opposed to aggregate-driven decisions that most of classical statistical decisions are. This approach makes it appear that data-driven decision making is non-statistical. The fact of the matter is that statistics is fundamentally the science of data, and must have data-driven decisions.

The literature on the decision theory is quite vast and is not restricted to the domain of statistics alone. It is a fact that Abraham Wald published his book ‘Statistical Decision Functions’ in 1950 and began the discussion on ‘deciding’ rather than ‘estimating’ as a statistical activity. One of the early collection of references on the decision theory is the book ‘Theories of Decision Making: An Annotated Bibliography’ by Anderson and Anderson (1977). The bibliography has five chapters and covers the four major perspectives of decision making, namely,

i) Rational decision making  
ii) Organizational decision making  
iii) Political decision making  
iv) Psychological decision making

The field of decision making has grown through the levels of (i) rationality, (ii) extensions of rationality, (iii) alternatives to rationality and (iv) a multi-perspective approach to decision making. Keeney (1982) describes decision analysis and explains what can and what cannot be done with decision analysis. The author then discusses why it is necessary to understand decision analysis as well as how it is done. In order to accomplish these purposes, the author begins with a discussion on the decision environment. The author then lists the following factors that contribute to the complexity of decision problems:
Multiple Objectives
Difficulty of identifying good alternative
Intangibles
Long-time horizons
Many impacted groups
Risk and uncertainty
Risks to life and limb
Interdisciplinary substance
Several Decision Makers
Value Trade offs
Risk attitude
Sequential nature of decisions.

Moreover the author highlights the following characteristics of today’s decision problems

High stakes
Complicated structure
No overall experts
Need to justify decisions

The essence of decision analysis can be summarized in the following

A perceived need to accomplish some objectives
Several alternatives, one of which must be selected
The consequences associated with alternatives are different
Uncertainty about the consequences
The possible consequences are not equally valued

Decision making methodology is goes through the following four steps:

Construction of the decision problem
Assessment of the potential impact of all alternatives
Deciding and determining the preferences given by the decision maker
Evaluation of and comparison and contrast among alternatives

The author also mentions the following misconceptions about decision analysis.

What is needed is an objective and value-free analysis, whereas what we have is subjective and value-based.
ii) Many decision makers do not fit in decision analysis because they violate the axioms of decision analysis by their choices of alternatives.

iii) Important factors are left out the analysis and hence the purpose of solving decision problems is hardly achieved.

iv) Most of decisions involve groups of decision makers whereas decision analysis requires a single and identifiable decision maker.

Axioms of decision analysis:

Decision analysis has an axiomatic foundation and the axioms are as follows.

Axiom 1. A: Generation of Alternatives
It is possible to specify at least two alternatives.

Axiom 1. B: Identification of Consequences
It is possible to identify potential consequences of every alternative.

Axiom 2: Quantification of Judgment
It is possible to specify the relative likelihood of every potential consequence that could result from an alternative.

Axiom 3: Quantification of Preference
It is possible to specify the utility of every potential consequence of any alternative.

Axiom 4.A: Comparison of Alternatives
When two alternatives result in the same two potential consequences, preference is given to the alternative which has a better chance for the preferred consequence.

Axiom 4.B: Transitivity of Preferences
If alternative ‘a’ is preferred to alternative ‘b’ and if alternative ‘b’ is preferred to alternative ‘c’, logically alternative ‘a’ is preferred to alternative ‘c’.

Axiom 4.C: Substitution of Consequences
If one consequence of an alternative is replaced by a set of consequences that have the same probability as the original consequence, then the original and modified consequences are considered to be equivalent.

1.2. Introduction to Decision Theory
To define decision theory we can say that it is no more than a theory about decision making in various contexts—simple as well as complex. It is not a complete and unified theory in
itself. There are many ways in which it has been understood and described by researchers. This chapter makes an attempt to summarize it while also reflecting on the diversity of the subject. Being an introduction, this chapter puts very little emphasis on technical aspects of decision theory. It may be appropriate to begin with some examples of decision problems that highlight the different aspects of decision making.

1. Should I take an umbrella with me?
The decision needs to be taken when I am not sure whether it will rain or not.

2. Should I buy this house?
When I look for the house to buy and when an offer is made to me, I need to take the decision whether I can find a still better house if I keep searching or whether I can lose the opportunity to purchase this house if I do not decide in time.

3. Should I smoke another cigarette?
One single cigarette may not cause a serious health problem, however if I make the same decision now and then, the decision may have an effect on my health.

4. Should the court declare the accused to be quality?
The court can make two mistakes, namely convicting an innocent person and acquitting a quality person. The question is about the principle that the court must apply if it considers the first of the two mistakes to be more serious.

5. What decision should a committee make if its members have different opinions?
The question is about what rules should be used to ensure that a decision will be reached even though the members are in disagreement.

Almost every human activity involves decisions. However, decision theory does not get involved in every human activity. It focuses on the available choices and on how the freedom to choose is used. In the situations considered by decision theory, choices are available and a non-random choice is made. In these problems, the available choices are directed toward a specific goal. Thus theory of decision is related with goal-directed behaviour in the company of numerous choices.

We should always keep in mind that decisions are not made in continuity. There are periods of decision-making and there are periods of implementing the decisions. Decision theory is concerned with the decision-making periods.

1.2.1 Decision Theory and its Interdisciplinary Nature
Modern decision theory was developed in the last sixty years. Several academic disciplines have, in one way or the other have significantly contributed to its development. Even though today it has developed as an independent academic discipline in itself, researchers working on decision theory include statisticians, economists, psychologists, political and social scientists, management scientists, and public administrators. There are some areas of specialization in these disciplines. For example, a political scientist is interested in the study of voting patterns and other aspects of collective decision making; a psychologist can be interested in studying the behavior of individual while making decision whereas a philosopher may be interacted in the requirement of rationality in making decisions. However, there is no watertight compartment between all these academic disciplines. Though they are likely to overlap one another, they are complementary to one another to large extent.

1.2.2 Normative and Descriptive Theories

In early days of the development of decision theory, there were two aspects, namely normative and descriptive theories. The theory about how decisions should be made is normative, while the theory about how decisions are made in reality is descriptive. It is interesting to note that there is no unique meaning of the word “should” in the description of normative theories. Nevertheless, decision theorists almost completely agree that it refers to the requirement that the resulting decisions are rational. It essentially means that normative decision theory develops norms for decision-making so that decisions are rational.

The sense of the world ‘normative’ however, is not so limited. It is also not implied that norms of rationality are the only norms that can be applied to decision-making. We understand that rationality is not the only important norms in making decisions. In addition to rationality, there are many other norms which need to be considered important in order to avoid ambiguity and unwarranted overlap of norms. According to this theory, one does not enter the domain of decision theory unless the norms that address ethical or political issues are already resolved. The normative theory addresses all normative issue after the goals have been identified. These issues typically consist of questions about how to take decision in presence of uncertainty and insufficient information. It also concerns issues of coordinating decision from, time to time and coordinating several individual decisions social decision procedures.

It is important to know the scope of decision theory by indentifying what is outside the consideration for decision theory. For instance, decision theory can provide a business executive
with different methods enabling him to maximize his profits. It can also provide an environmental protection agency with different methods which are useful to it in order to minimize hazardous exposure. However, the theory does not answer the basic question whether these targets should be attempted to attain. Though the scope of the ‘normative’ part of decision theory is very limited, normative and descriptive interpretations decision theories are vague and ambiguous. Finding examples of a glorying ambiguity and even confusion between normative and descriptive interpretations of the same theory has been a very common practice. What is more important, nevertheless, is to recognize the fact that it is more difficult to draw sharp boundaries between normative and descriptive interpretations in decision making as a discipline and also in many other academic disciplines. This point can be illustrated through consideration of falsification of a decision theory. The parameters of the falsification of an expressive decision theory, in particular, are quite obvious. There are many parameters or criteria used to consider falsification of a decision theory. The most popular of these criteria are the following:

Criterion I: If a decision making theory finds a decision problem where most people act in contradiction to the theory, it is falsified as a descriptive theory. A normative theory leads towards a rational decision making. Falsification, therefore, is required to refer to the necessity of rationality in decision making. However, the nature of the conflict between the theory and actual decision making in day today life is not crystal clear. It is therefore proposed to have two criteria according to the strength of the conflict.

Criterion II: If we find a problem in decision making where the decision maker makes any decision without being rational, the decision theory is likely to be falsified as normative theory.

Criterion III: A decision theory is believed to be strongly untrue as a normative theory if we discover a decision problem where the decision maker acts according to the theory but not rationally.

To cite an interesting example, let us consider a theory X which is claimed as valid as a normative as well as a descriptive theory by its developer. If we know a decision problem Y where most of the decision-makers do not comply with X, This implies that X satisfies criterion I. However, what happens in practice is that many of the decision-makers claim that the theory X satisfies criterion II, or even criterion III. It suggests that decision making is accompanied by the claims of its invalidity from a normative point of view. This happens frequently in real life decision-making problems and can be confusing for the analyst.
1.3 Decision Processes

Most of the decisions take time and cannot be made instantaneously. Decision processes are therefore divided in stages or phases. Some of these concepts are described here.

1.3.1 Condorcet

The great philosopher Condorcet (1743-1794) has put forward a general decision making theory comprising three stages of the process of a decision making. The first stage, according to him, involves discussing the principles serving as the basis for decision making and examining different aspects of this matter and possible consequences of traditional decision-making. This phase involves only personal, opinions and there is no attempt to form a majority. The second stage that follows the first involves a discussion for clarifying questions, approaching and combining opinions to form a small number of opinions that are of a more general nature. This reduces the decision to choose among a possible set of choices or alternatives. The third stage, according to him, constitutes the actual decision by choosing from among the available alternatives.

Condorcet’s theory is very insightful, but his theory of the stage in a decision process seems to have been forgotten. More details of the theory and a good discussion can be found in Condorcet (1793). Condorcet himself described the three stages of a decision process as discussion, clarification and choice. He also suggested that decisions are temporally bounded.

1.3.2 Modern Sequential Models

John Dewey (1978) ignited interest in a sequential model for a decision process by proposing a five stage model instead of the Condorcet model of three stages. The five stages of the Dewey model are as follows:

Step I: Identifying and defining the problem
Step II: Analyzing the problem and establishing the criterion
Step III: Generating creative and productive solutions
Step IV: Evaluating the options and selecting the best possible solutions
Step V: Taking action leading to the required solutions

Herbert Simon (1960) made certain changes in Dewey’s list of six stages and modified them so as to make them appropriate for organizational decision making processes. Simon propagated a three-phase decision-making process, where the three phases are:
Phase I: Intelligence, that is, finding occasions for making a decision
Phase II: Design consisting of possible courses of action
Phase III: Choice involving selection from the available courses of action

Brim and others proposed yet another division of a decision making process comprising five steps in 1962. The stages he proposed are as follows:

I. Identifying the problem
II. Acquiring necessary information to solve the problem
III. Generating probable and possible solutions
IV. Evaluating solutions generated
V. Selecting a strategy to act and perform accordingly

They have also added the sixth stage of implementation of the selected decision to solve the problem.

The decision making theories and their stages proposed by Dewey, Simon, and Brim et al. are ordered in a particular sequence, i.e. these theories are sequential in nature. Witte (1972) and several other researchers have criticized the concept that a decision process is generally to be divided into successive stages. As a matter of fact, Witte (1972) has concluded as follows:

We judge that human beings cannot collect information lacking in some way at the same time rising alternatives. They cannot keep away from evaluating these alternatives without delay, and in doing this they are mandatory to a decision. It is a package of operations and the series of these packages over time constitutes the total decision making process.

This discussion leads to the conclusion that a realistic decision model must allow different stages of the process of decision making to occur in different orders and different contexts.

1.3.3 Non-Sequential Decision Models

According to Mintzberg, Raisinghani, and Theoret, (1976), a decision process goes through different stages, but these stages are required not to have sequential relationships among them. The three stages they proposed are similar to those proposed by Simon. However, these stages have been renamed as identification, development and selection.

The stage one of identification is known as intelligence by Simon. It, in his opinion, contains two routines – decision recognition in which the problems and opportunities are identified as soon as a decision-maker receives a stream of ambiguous and mostly verbal data,
and ‘diagnosis’, where existing information channels are tapped and new ones are opened in order to define and clear the issues.

The development stage is called ‘design’ by Simon. It, according to him, serves to define and explain the existing options. This stage also contains two routines - ‘search’ and ‘design’. The object of the first routine called search routine is to discover ready–made solutions to the problems while the second routine known as design routine intends to modify read-made solutions or to develop new solutions to the problems.

The selection stage, known as ‘Choice’ stage, to use Simon’s term, has three routines. The first routine is called of a ‘Screen’ routine. A screen routine is required only if a screen stage generates additional handy materials than can been intensively evaluated. The options or alternatives which are clearly suboptimal are naturally eliminated in the screen routine. The second routine, known as evaluation choice routine, involves the actual choice from the obtainable alternatives. This routine is likely to make use of make use of three possible modes of comparison - judgment, negotiation and analysis. The third routine, known as ‘authorization’, expects to acquire sanction of the solution chosen at the higher stage in hierarchy.

All these phases and routines are related more circularly than linearly. In this scheme of decision making it is possible that the decision-maker moves within the limits of identification in order to the nature of the issue during the design, moves through a web of complicated designs and search activities in an attempt to arrive at an expected solution during the course of evaluation, and alternatively continues his search stuck between development and investigation in order to appreciate the nature of the problem in hand. It is also possible that the decision-maker moves between selection and development of existing alternatives with the only purpose of reaching his goal. If no acceptable solution is found, the decision-maker moves back to the development phase. For more details, see Mintzberg et al. (1976).

1.3.4 Phases of a Practical Decision Theory

Simon (1960) has observed that executives spend a very limited time in making a choice; they use a huge portion of their time in actions related to intelligence, and even a larger portion of time in design activity. Mintzberget al. (1976) have elaborated this through their empirical
findings. They have proved that the development stage dominates the other two stages with the help of their study of 21 out of 25 decision processes.

This observation is contradictory to the history of literature on decision-making which is, to a large extent devoted to the study of evaluation-choice routine. Although many of the so-called empirical approaches take the entire process of decision process into consideration, decision theory appears to be solely concerned with the evaluation-choice routine. Mintzberg et al. are curious because this habit seems to be faraway less important in many of the decision processes, we have studied than diagnosis or design. (Minizberg et al., 1976).

What can be said in the defense of the decision theory is, that the evaluation option routine is at the middle of the whole decision process. This routine is what constitutes the process of decision making and therefore the features of the other routines are mostly determined by it. Therefore, it is necessary to pay a great deal of concentration especially on the evaluation choice routine. However, we should not neglect other routines completely. This exactly is the fault of normative decision theory in most cases.

1.4 Values and Decisions

While making decisions or choosing among given options, the attempt needs to be made to obtain a result which is as good as possible according to some standards of deciding what is good and what is bad. It is believed that the selection of a value or a standard for decision making is more philosophical than mathematical. Decision theory takes it for granted that such a value or standard is available to define the process in an unambiguous and useful way.

1.4.1 Numbers and Relations

There are two basic ways of expressing the standard or the value. One way is to express it as a relation between choices. For example, choice A is better than choice B. There is another way of expressing the value or the standard. Numerical values can be assigned to all the available choices and alternatives. This is nothing but numerical representation of the value pattern. If a higher value represents a better choice than the best decision it is better to select the option which has the highest numerical value. This proves that the two common ways of expressing the value pattern are relational and numerical according to which a decision can be made.

1.4.2 Comparative Values
There are two different ways of representing values - relational representation of values and numerical representation of values. If we compare these two ways of representing values, relational representation of values is very common in decision making in day-to-day practice. Alternatives are compared using Phrases like ‘as good as’, ‘as bad as’, ‘better than’, ‘worse than’, ‘equally good’, ‘equally bad’ and so on are mostly used while comparing existing choices or alternatives. As two alternatives are compared and related with each other, the relations between two alternatives are called a binary relationship. In order to put it in simple words, the mathematical form, A > B in its place of the verbal explanation A is better than B is generally used. However, it should always be kept in mind that, the terms ‘betterness’ or ‘worseness’ are not symmetrical when used in everyday life. To say that A is better than B is not similar to B is worse than A. There can also be psychological asymmetries (Tyson, 1986 and Houston et al., 1989). Though, as the difference between ‘not worse’ and ‘better’ is not sufficiently significant and is ignored in decision theory. As a result, ‘A > B’ is interpreted as B is worse than A as well as A is better than B. There is yet another term used in the value comparison and the term is of equal value. The notation used to represent this relation is ‘≡’, and ‘A ≡ B’ meaning ‘A’ and ‘B’ have the same value according to the standards being used. Finally, there is a phrase ‘at least as good as’ which is denoted by the symbol ‘≥’. These three comparative relations, namely better than (>), equal in value to (≡), and at least as good as (≥) are the basic components of an official language of preference logic. In this preference logic, > represents preference or strong preference; ≥ indicates weak preference; and ≡ indicates indifference. The following two rules are considered to show the connections among these three relations:

I. A > B if and only if A ≥ B and not B ≥ A

II. A ≡ B if and only if A ≥ B and B ≥ A

A vast literature can be found on the mathematical properties of these three relations. Two of these properties are often referred to in the context of decision theory and hence are discussed here. These are completeness and transitivity.

**1.4.3 Completeness**

Every preference relation signifies a group of entities with reference, which it’s defined. This set is called the domain of the relation. Completeness is referred for a relation and its domain.
Completeness: ‘A relation ≥ is said to be complete only if, for any two elements A and B of its domain, either A ≥ B or B ≥ A.’

Completeness is also known as connectedness.

It is a common experience that the preference relation being used is not complete. Nevertheless, preference completeness has an academic value and it is generally accepted as a simplifying and clarifying assumption. It has been also a common practice to assume preference completeness, even though it can often be a questionable assumption.

1.4.4 Transitivity

‘A strict preference relation > is transitive if it holds for all elements A, B and C of its domain and if A > B and B > C ⇒ A > C.’

In the decision theory, it is taken for granted that strict preference (that is, >) as well as weak preference (that is, ≥) and indifference (that is, ≡) are transitive.

Like completeness, transitivity generally is a problematic assumption in decision theory.

1.4.5 Use of Preferences in Decision-Making

In decision-making an attempt is made to use preference relations in order to choose the best alternative. This purpose may be served by the following simple rule.

Rule 1: Alternative is best, if it is better than all other alternatives. As soon as we find the best alternative, it is selected.

There can be a situation when there is no uniquely best alternative because there are two or more alternatives that are “best” by the foregoing definition. In such situations, an obvious solution is to select any of the “best” alternatives. The following rule makes it explicit.

Rule 2: Alternative is the best and not necessarily unique if, it is at least as good as all other choices and or alternatives. If 2 or more than two alternatives are equally best, we need to select any one of them.

It is to be kept in mind that preference relations that do not satisfy rationality criteria like transitivity are most often not useful in decision-making.

1.4.6 Numerical Representation

It is possible to use numbers to signify values of the choices which are to be subjected to the decision process. Generally speaking, a relational representation can be obtained when there is a numerical value assignment. Moreover, the weak preference relation obtained in this way is
always complete in itself and all other preference relations like strict and weak preference and indifference are likely to be transitive.

We may encounter certain problems in this approach. One possible problem is that it may not always be clear what the numbers represent. Assigning numbers to represent relation may appear to be arbitrary if there is no measure of goodness represented by numbers.

Some moral theorists have claimed that it is possible to reduce all values to a single entity known as utility. This entity may or may not represent human happiness. However, the utilitarian moral theory perceives that all moral decision attempt to maximize the total amount of utility, at least in principle. The utilitarian theory, in this way, gives a new decision theory which is based on the use of mathematical representation of the value.

1.4.7 Use of Utilities in Decision-Making

It is easy to use numerically represented values, or utilities in making of decisions. The essential decision rule is simple, lucid, precise, and clear.

Utilitarian Rule 1: Select the alternative that has the highest utility.

This rule, however, is not applicable if a decision maker finds more than two choices or alternatives almost with the same maximal utility. In such cases, the rule needs to be modified as follows:

Utilitarian Rule 2: Select the alternative that has the highest utility. If two or more alternatives have the highest utility, select any one of them.

This is the rule of maximization. Some critics of the utilitarian theory maintain that this is too much to expect. It is more reasonable to select alternatives which have the levels of utility less than the maximal utility, but at the same time which are acceptable. Decision theory almost universally employs the maximization rule.

1.5 Standard Representation of Decisions

The next important development in decision theory is the introduction of decision matrices, which have become the standard representation of a decision problem. This involves some more basic concepts of decision theory, namely alternatives, outcomes and states of nature.

1.5.1 Alternatives

A decision consists of selecting from different alternatives providing tricky options. These alternatives are the possible courses of action available to the decision-maker during the time of making a decision. A set of alternatives can be defined precisely and reasonably. In a particular
decision problem, it is said to be open if new alternatives can be added to it. In other decision problems, it is said to be closed because no new alternatives can be added. There is a limited and fixed number of alternatives to choose from. A closed set of alternatives may be divided into two subsets, a set of voluntary closure and a set of involuntary closure. The voluntary closure is the set that is closed by the decision-maker, possibly as the first step in decision-making. Closure is imposed on decision maker by external forces or circumstances in the case of involuntary closure.

Open alternative sets are extremely common in day today life, but alternatives sets are usually assumed to be closed in a decision theory. This is so because theoretical treatment finds decision problems accessible under the assumption of closure. A definitive solution would not generally be available if the alternative set is open. Moreover, it is also a common practice to assume that alternatives are mutually exclusive. This implies that it is not possible to realize two of the alternatives at the same time. Put together, it is important to note that the decision theory assumes the set of alternative to be closed and its elements mutually exclusive.

1.5.2 Nature and Outcomes

A decision maker’s choice of an alternative out of existing alternatives and the way he carries it out determines the outcome of a decision. The effect of decisions also depends on some factors which are not under the control of a decision-maker. Some of these external forces arise out of information available in the background, and this background information is available to the decision maker. There can still be some factors that depend on decisions of other people and on certain traits of nature which are unknown to the decision-maker. It is a common practice in decision theory to aggregate the different unknown external factors into a set of cases and label them as states of nature.

The possible outcomes of a decision are then easily seen to be the result of the selected alternative and the prevailing state of nature. Since every outcome results from the set of various alternatives with their particular state of nature, it is common practice to present the situation in the form of a decision matrix.

1.5.3 A Decision Matrix

A decision matrix provides a standard set-up for the assessment choice routine in theory of decisions. The alternatives available to the decision maker are experienced vs all the probable states of nature in the decision matrix. The rows in matrix represent the alternatives available
whereas the columns represent the states of nature. The decision matrix assigns an outcome for each and every alternative available and every possible state of nature. Decision analysis using the decision matrix additionally requires the following:

1) Information about the method of evaluating the outcomes, and
2) Information about possible realization of the states of nature.

Assigning utilities to outcomes reached is, perhaps, one of the most common ways of representing the values of the outcomes. Then it is possible to make use of utility values instead of verbal descriptions in the matrix.

Decision theory exclusively handles only those problems which can be represented in the form of such matrices, the utility matrices. Most of the current decision-theoretic methods demand numerical information. It may not always be possible to have precise value information. However, it must be kept in mind that constructing methods that can deal with non-numeric information is much more difficult.

1.5.4 The State of Nature

Decision theory combines in itself utility matrices in the company of different types of information available about the state of nature. In an extreme situation, the decision maker may know which state of nature will occur. In such a case, only one state of nature needs to be considered. Such a case is described as making decision under certainty. If you know the outcome for every alternative you can choose, then you are acting under certainty. Otherwise you are acting under non-certainty.

It is a common practice to divide non-certainty in categories like risk, uncertainty and ignorance. This subdivision is given by Knight who categorically highlighted that the term ‘risk’ covers two things which are categorically different in 1921. One meaning of risk is ‘a quantity susceptible of measurement’, while the other meaning of risk is something clearly not of this character. Knight suggested to use the term uncertainty to refer to the case of non-quantifiable form, and the concept risk to refer to the quantifiable cases. The literature has the following definitions of the three terms:

The decision-making is being done under

1. Certainty: When every action invariably has a specific outcome.
2. Risk: When every action has a set specific possible outcome and each possible outcome can occur with a known probability. It should be noted that, certainty is a unique case of risk for which the probabilities are either 0 or 1.

3. Uncertainty: When every problem has a set of exact outcomes as its result, the probabilities of these outcomes are not known. (Luce and Raiffa, 1957). It may be worth noting that the three alternative listed above are not mutually exclusive. As a matter of fact, most of the decision problems involve risk and uncertainty. It is has been also a standard practice to employ the term uncertainty in order to coat the situations with only a limited knowledge of probabilities. Luce and Raiffa called a strict uncertainty as called ‘ignorance’. (Alexander, 1975). All the previous discussion can be summarized in the statement that an accepted representation of any decision comprises a utility matrix as well as some information on the degree of uncertainty that occurs. Hence a probabilistic assignment to every state of nature is included in the standard representation of a decision making situation under risk.

1.6 Expected Utility

Expected utility dominates all approaches to decision-making under risk, that is, when probabilities are known Shoemaker (1982) states that it is the main paradigm in decision making as the second world war.

1.6.1 Defining Expected Utility

A precise description of the expected utility is ‘probability-weighted utility theory’ which assigns a weighted average of the utility values in the context of different states of nature where the probabilities of these states are used as weights. Suppose there are n outcomes, each having associated utility and probability. The utilities are denoted by u₁, u₂, …, uₙ while the probabilities are denoted by p₁, p₂, …, pₙ. The expected utility is then given by

\[ u ≡ p₁ u₁ + p₂ u₂ + … + pₙ uₙ \]

The theory of expected utility was developed during the 17th century along with the probability theory in the context of parlor games. Arnauld and Nicole (1662) have stated that to decide what one have to do to get a good or to get out of an evil, one must not only consider the good and the evil in itself, but also the probability that it will or will not happen and view geometrically the proportion that every things have jointly.

1.6.2 Subjective and Objective Utility
Expected utility theory refers to monetary outcomes in its early days. It recommends participating in a game if it increases the expected assets. The probabilities assigned to results are objective. Nicolas Bernoulli (1713) poses a problem in probability theory that is now popular by the name ‘St. Petersburg paradox’. It involves the following game: An unbiased coin is tossed until the result is the first head. For a positive integer \( n \), if the first head occurs on the \( n \)th toss, then the player wins \( 2^n \) gold coins. The probability of getting the first head on the \( n \)th toss is \( \frac{1}{2^n} \) and therefore the expected amount won after the game is over is:

\[
\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \cdots + \frac{1}{2^n} \times 2^{n-1} + \cdots
\]

Since the series does not converge to a finite limit, the expected wealth from the game is infinity. Therefore, according to the principle of achieving the expected utility, any rational person is willing to pay any amount, as long as it is finite, for playing this game. Daniel Bernoulli (1738) proposed a solution to the St. Petersburg puzzle. He began with the idea of instead of maximizing the expected utility, instead of maximizing the expected wealth. He then argued that utility that a person attaches to wealth does not increase linearly with the amount of wealth, but increase at a decreasing rate. It is then quite easy to verify the fact that any human being with such a utility function would not be willing to play the game. Subjective utilities are commonly used while applying this decision theory to solve economic problems. On the other hand, objective utility approach dominates risk analysis. Bondi (1985) has concluded that the recommended way of measuring risk is to grow the probability of a risk by its strictness, to describe this as the expectation and then to use this expectation to evaluate risks.

1.6.3 An Assessment of Expected Utility

The principle of maximizing expected utility is most commonly favored because it is reasonably safe in the long run. The principle of maximizing expected utility is not valid for all decisions where the involved events are unique or very rare. Nevertheless, the very small probability of the occurrence of such an event does not affect the general principle. As a matter of fact, the leveling-out effect dominates more as the set of decision that are covered by the rule grows larger. More specifically speaking, large catastrophic outcomes get leveled out if the set of decisions is sufficient large. On the other side of the argument, it is also true that the challenges of information processing may lead to disproportionate losses if too many issues are put together. It is useful to note that decisions can be divided into certain manageable parts in one way or the
other. The way it is done can influence the decision outcomes. Even the cases where it is valid to argue for expected utility maximization having the leveling-out effect do not call for rationality. To be more specific, it is perfectly wise and rational to abstain from minimizing total harms so as to avoid high probability risks.

It is therefore necessary to note that the principle of maximizing expected utility is meaningful only when the comparison is between options in a single decision. Risk analysis may involve a clear violation of this basic requirement. It is often found that expected utility calculation may be used for comparing risk factors, for which no options in one and the same decision. It can then be said that the normative nature of expected utility maximization depends on the expected leveling-out effect. The major argument in its favor is in cases where a large number of decisions are made according to one and the same decision rule.

1.6.4 Estimation of probabilities

Calculation of expected values requires reasonably accurate estimates of probabilities of outcomes. Some application use empirical frequencies for this purpose. Most of the cases, however, do not have a reliable basis for estimating these probabilities. The reliability of probability estimates depends on dissimilarities between objective probabilities and their subjective estimates. Such dissimilarities are explained by short of calibration in psychological experiments. The estimates of probability are likely to be calibrated if the true proportion equals the assigned probability in the long run. (Lichtenstein et al., 1982). Experimental studies have found out that only a small types of estimations are made in a calibrated approach by experts. Most of the predictions have been found to be subject to over confidence. It is proposed by Lichtenstein et al. (1982) that the effects of over assurance in estimating probabilities may be very serious. As a result, subjective estimates of (objective) probabilities are often not sufficiently reliable. Then, the principle of maximizing expected utility cannot find a strong support if the estimates of probability values are only subjective.

1.7. Generalization of Expected Utility

Large numbers of models used for making the decisions under risk are generalizations of expected utility theory. The following two of them are sufficiently important to discuss.

1.7.1 Process Utility and Regret Theory

Expected utility assesses option in terms of the utility of each outcome irrespective of other possible outcomes. A decision-maker can be influenced by other considerations like a
desire to avoid uncertainty and the relations among the outcome reached and all other probable outcomes. Such values are called process utilities (see, Sowden, 1984) and are represented by numerical values. Even though expected utility theory does not allow process utilities, it is reasonable to value certainty, as such and there is no argument that refutes this presumption (see Sowden, 1984). A general expected utility theory takes process utilities into reflection, and permit, for the impact of attitudes on risks and certainty. Weirich writes: “It makes expected utilities sensitive to risks by including risks inconsequence”(1986).

There have been many arguments against the generalized expected utility. Most arguments are likely to be valid for the generalized expected utility theory in the most universal appearance if there are no boundaries, on the number and, types of process utilities. Nevertheless, these limitations and restrictions may be important to construct the theory falsifiable with not losing its important advantages. The theory that achieves this is called regret theory. In this case, regret means the undesirable recognition of the fact that ‘What is’ is more unfavorable than ‘what could be’. The converse of this comparison is called ‘rejoicing’. Regret theory uses a two-attribute utility function which is based on two measures of satisfaction - utility of outcomes and quantity of regret (Loomes, and Sugden (1982), Bell (1982) and Sugden (1986).

In its simplest form, Regret theory measures regret in terms of the difference among the actually received assets and the highest level of assets formed by additional alternatives. The utility function is written as \( u(x, y) \), in which \( x \) denotes the actually received assets and \( y \) denotes the difference previously mentioned. It is found reasonable to consider \( u(x, y) \) to be monotone increasing in both \( x \) and \( y \).

Regret theory can explain how a person can gamble (indicating risk-prone decision) and purchase insurance (indicating behavior averse to risk) at the same time. Both these behaviors can be explained in terms of attempts to avoid regret, rather than minimize risk.

1.7.2 Prospect Theory

Kahneman and Tversky (1979) have proposed and developed the prospect theory in order to explain the results of experiments in the contexts in which decision problems are described in relation with monetary results and objective probabilities or possibilities. It is also relevant to a general process of decision-making. Prospect theory, thus, differs from many other theories on decision-making because it is blatantly descriptive and makes no normative claims. It is unique
in the sense that it differentiates between the editing stage evaluation stage involved during the process of decision making. The editing stage organizes and reformulates the options in an attempt to simplify the nature of evaluation and the selection of choice. It also identifies gains and losses and defines them in the context of a reference point which is neutral. The reference point, to a large extent, is likely to correspond to the existing position of the asset. However, it can be influenced by the prospects, offers, requirements and expectations on the part of a decision-maker.

The evaluation phase, on the contrary, evaluates the options edited during the editing phase. Prospect theory proposes this evaluation to be done as if two scales are available to the decision-maker. One takes place of the monetary outcomes while the other takes place of the objective probabilities from the original problem. Monetary outcomes are substituted by a value function ‘V’ which assigns a number V(x) to each outcome x, so as to reflect the personal and subjective value of the concerned result and outcome. Generally speaking, the value function is no more than the function of monetary outcomes used to measure the subjective utility. The main characteristic of this value function is that it is applied to changes rather than to final states. The value function differs from different reference points and therefore is considered to be the function of two arguments written as V(w, x), in which w is the present state of wealth. The significance of a prospect theory lies in the fact that it transforms objective probabilities with the help of a function $\pi$, known as the decision weight. Its function increases from [0,1] to [0,1]. It is used in place of probabilities especially in the utility theory. However, it fails to follow and satisfy the laws of probability. It is also explained as something used to measure the degree of belief. Two important properties of the decision weight function are worth mentioning. First, differences among probabilities close to certainty are over weighted. Second, the weight function is not defined for probabilities that are close to zero and one.

Although the prospect theory makes no claim of being normative, it gives two lessons to a normative theory. The first lesson is that the editing phase and framing the decision problem is very important. It is important to focus on the rational expectations and demands during the course of framing the decision problem. The second lesson is that the tendency of either ignoring or over weighting small probabilities has important normative aspects.

1.8 Decision-making in Uncertain Contexts

1.8.1. Paradoxes of Uncertainty
There are two paradoxes of uncertainty which are at the centre of the discussion on the difference between uncertainty and probability. One of the paradoxes of uncertainty, discovered by Pierce (1878), is known as the paradox of perfect evidence. However, that paradox was formulated by Popper (1959). This paradox is formulated as follows:

Suppose 'c' is an unbiased coin, and let 'b' be the statement that the n\textsuperscript{th} toss of 'c' will give up heads. In the subjective theory, the probability of statement 'b' may be assumed to equal to ½. That is,

\[ P(b) = \frac{1}{2} \quad \ldots \ldots \quad (I) \]

Let 'e' be some evidence, possibly in the form of observation of several tosses of 'c', and let 'e' be preferably favourable to the hypothesis that 'c' is firmly symmetrical. Then we have

\[ P(b,e) = \frac{1}{2} \quad \ldots \ldots \quad (II) \]

This proves that the probability of tossing heads remains same, even in the light of the evidence 'e'. We, therefore, have

\[ P(b) = P(b,e) \quad \ldots \ldots \quad (III) \]

Though, the subjective theory interprets (III) as “e, on the whole, is irrelevant with respect to b”. In other words, this is like saying that our supposedly rational belief in the hypothesis would be totally not affected by the accumulated evidential information ‘e’.

Pierce suggested that not one but two number are required to express the proper state of belief. The first number depends upon the inferred probability whereas the second number depends upon the information on which the probability is based.

The second paradox is Ellsberge’s Paradox and relates to the following decision problem:

Consider the box containing 30 red balls and 60 balls that may be black or yellow and the exact number of black or yellow balls is not known. A ball is to be selected, at random and the player has to bet on the colour of the ball. Four actions are specified below for comparison:

The first action is the bet on red.
The second action is the bet on black.
The third action is the bet on yellow or red.
The fourth action is the bet on yellow or black.

The most frequent patterns of responses are found to prefer first action to these second, and fourth action to the third.
The paradox arises because these preferences are inconsistent with the predictions of expected utility theory. More information can be obtained from Ellsberg (1961).

1.8.2. Measures of Probabilities Not Completely Known

Almost all the rules that have been stated and developed so far for good decision-making under uncertain context use some quantitative expressions of partial probability information. These are called measures of uncertainty and some of these are discussed here. Basically there are two types of measures that measures not completely known probabilities. These are traditionally called binary and multi valued measures. The binary measure is the measure that classifies the probability values into two groups - the possible values and impossible values. It has been a common experience that a set of possible probability values constitutes an interval. Ellsberg (1961) used these binary measures and referred to them as a set of sensible probability judgments. Levi (1986) refers to them as a group of permissible chance judgments.

Multi-level measures are usually represented by a function which assigns a arithmetic value, foreach and every probability value. This value symbolizes the reliability of the concerned probability value. Literature gives different interpretations of this measure some of which are as follows:

1. Second Order Probability

The reliability measure indicates that the true probability has a specified value and significance. This can be considered as a subjective probability as against and objective probability which, too, has a certain value. It can be interpreted that the prejudiced probability might have taken a sure value if there is approach to certain specific amount of information. However, the fear of an infinite regression of higher orders of probability has resulted in a negative attitude among philosophers and statisticians. Hansson (1975) has stated that “merely adding second order probabilities is not a real solution because we cannot be certain about these (that is, second-order) probabilities.” Even though this is not the place to discuss second-order probabilities, it is mentioned only to highlight the fact that identical arguments can be devised for other measures of incomplete probability information. An accurate formalization is necessary, for the precision, in a probability estimate.

2. Epistemic Reliability
P is a real valued measure which lies between 0 and 1, is assigned to each probability and is called the ‘epistemic reliability’ of the concerned probability. (Gardenfors and Sahlin, 1982) No mathematical properties are specified for ‘P’.

A binary measure derived from a multi-valued measure. Suppose ‘M₁’ is a multi-valued measure, then a binary measure ‘M₂’ can be given as follows:

For a real number ‘r’, define \( M₂(p) = 1 \) if \( M₁(p) \geq r \) and 0 otherwise. A multi-valued measure contains more information than a binary measure. On the other hand, it is simpler to express uncertainty since an interval somewhat than as a real valued function.

1.8.3. Decision Criteria for Uncertainty

Literature has provided ample decision criteria for decision-making under uncertainty. Five of these decision criteria are discussed below:

1. Maximin Expected Utility (MMEU):

   This rule requires choosing the alternative such that its lowest possible expected utility is as high as possible. It is treated as a decision criterion which is pessimistic.

2. Reliability-weighted Expected Utility:

   Each probability is given a weight attached by its degree of reliability and the weighted average of probabilities is calculated in the case of multi-valued decision measure. This weighted average is used for calculating an expected value for each alternative. This rule is known as reliability-weighted expected utility. It is treated as unduly optimistic as against pessimistic decision criterion.

3. Ellsberg’s Index

   Daniel Ellsberg (1961) suggested a combination of maximin expected utility and reliability weighted expected utility. He has assumed a group \( y₀ \) of possible probability distribution and a single probability distribution \( y₀ \) representing the best probability estimate.

   Let \( P \) denote the degree of confidence in \( y₀ \), let \( \text{min} \ x \) be the minimum expected pay-off for action \( x \), and let \( \text{est} \ x \) be the expected pay-off for action \( x \). The decision rule is to associate the index

   \[
   P \times \text{est} \ x + (1 - P) \times \text{min} \ x
   \]

   With each \( x \), and select \( x \) that has the highest index value. Here \( P \) is used to minimize the degree of optimism or pessimism as favored by the decisionmakers.

4. Modified MMEU
This rule used a measure ‘P’ of epistemic reliability over a set of probabilities. A certain level, i.e. \( P_0 \) is chosen for epistemic reliability. Probability distributions with lower reliability as compared to \( P_0 \) are deliberately excluded from any considerations because they are not considered to be serious possibilities. The MMEU criterion is useful to a set of probability distributions which have serious possibilities.

This above rule has two extreme cases. The first case signifies that if each and every one probability distributions possess identical epistemic reliability, the law gets reduced to the maximin rule. The second case implies that when only one probability distribution possesses non-zero epistemic reliability, the rule becomes Bayesian.

5. Levi’s Lexicographical Test

Levi (1973) takes the existence of a collection of permissible probability distributions and a collection of permissible utility functions for granted. A series of lexicographically ordered tests are proposed for decision-making under uncertainty. These three tests can be treated as three successive filters. The first is the test of E-admissibility. An option is E-admissible if the best of some combination from a permissible probability distribution and a permissible utility function. The second test is P-admissibility and it is applied only E-admissible options. An option is P-admissible if it is the most excellent with regard to the keeping of E-admissible options. The third test is S-optimality. A P-admissible option is S-admissible if it is ‘security optimal’ with regard to a number of permissible utility function. Security optimality approximately corresponds to the MMEU rule.

The maximin rule is pessimistic but, when applied to E-admissible and P-admissible options, it does not produce as pessimistic criterion.

1.9. Decision-Making under Ignorance

Under ignorance the decision making is no more than making decisions when the possible states of nature are known but their probabilities are unknown. This is the case of classical ignorance. A more severe case of ignorance is when even all possible states are not known. This case is not discussed here as it is beyond the scope of the study.

1.9.1. Decision Rule for Classical Ignorance

The most popular criterion suggested for decision making under ignorance is the maximin rule. Under this rule, the security level is determined for each alternative as the worst probable result for that option. The maximin rule then selects the alternative that which has the
maximal security level. The maximin rule, thus, maximizes the minimal result. VonNeumann first proposed this rule against an intelligent opponent. Wald (1950) extended it against nature. Since the maximin rule treats all alternatives having the security level equally, a variant is developed that compares second-worst outcomes when the worst outcomes contain the identical security level. When the second worst and the worst outcomes have the equal security levels, the third-worst outcomes are compared and contrasted. This continues until all ties are broken. (see Sen, 1970). This rule is known as lexicographic maximin or leximin rule.

The maximin and leximin rules are considered to be pessimistic. The other extreme is a result of the maximax rule which selects the outcomes that has the best hope level (that is best possible outcome). Rapoport (1989) states that, in general, it is not easy to give good reason for the maximax principle, as a rational principle because it reflects wishful thinking. It is obvious that a decision rule should not go to either of the two extremes, namely pessimism and optimism. This is possible only if utility information is available. Hurwicz (1951) proposed a middle way in the form of the optimism-pessimism index. This criterion requires the decision-maker to choose an index \( \alpha \) between 0 and 1. Suppose min (A) is the security level of the alternative A and max (A) is its hope level. The \( \alpha \)-index of A is then given by

\[
\alpha \times \min (A) + (1 - \alpha) \max (A)
\]

Utility information also provides the minimax regret criterion introduced and proposed by Savage. This rule recommends choosing the option that has the lowest maximal regret.

To sum up the above discussion, we can say that pessimism is a unique feature of maximin and leximin rules which make use of preferences as requirement information whereas optimism is a unique feature of maximax rule which also makes use of preferences as required information. \( \alpha \) - index rule which makes use of utilities as required information has its own feature, i.e. varying with index. Minimax regret rule, which, too, makes use of utilities as required information has a different feature, i.e. cautiousness.

Statistical methodology basically aims at drawing inference about the unknown parameter (also known as the state of nature) on the basis of observed data since the state of nature is unknown, statistical inference can be described as acting under ignorance. By contrast, decision theory deals with situations where consequences are observed after a decision is made. Since these consequences cannot be controlled by the decision alone, there is uncertainty about the consequence that actually occurs as the result of a particular decision. This situation therefore
can be described as acting under uncertainty. While ignorance refers to lack of knowledge about the present (that, existing but unobserved or possibly unobservable) state of nature, uncertainty refers to the non-deterministic nature of the future events that have not taken place. Both ignorance and uncertainty involve some risk of leaching to unfavorable consequences. It is therefore a common practice to put these two situations together under the label of “decision making under risk”. Mathematically speaking, risk is defined as the expected loss. This is the reason why risk theory involves probability theory. Never the less, it is necessary to distinguish between inference and decision. Probability theory is objective in the sense that it does not directly provide critical levels for changing the decisions. The choice of critical levels depends on value judgments and also on probabilities in some way. By assessing the risk, the question being answered is what are the consequences of making incorrect decisions? and what are the costs of carrying out more tests? Statistical inference in some sense is the conclusion arrived at after carrying out a statistical analysis of the available data. On the other hand, a decision must indicate a definite course of action.

Statistical inference, in this sense, is objective because it does not go beyond data. However, decision making involves much more than an objective assessment of the expected profit that may be incurred as the result of decision. Consider an apparently simple situation described as follows. A gambling house gives three options to the gamblers and every gambler must choose one of the three options and play according to the choice. The three options are as follows:

Option 1. The gambler receives Rs. 50 from the gambling house.
Option 2. An unbiased coin is tossed and if it turns up heads, the gambler receives Rs. 100 from the gambling house. If the coin turns up tails, the gambler loses, and neither receives anything nor pays anything.
Option 3. An unbiased coin is tossed and if it turns up heads, the gambler receives Rs. 125 from the gambling house. On the other hand, if the coin turns up tails, then the gambler loses and must pay the gambling house a penalty of Rs. 25.

Which of the three options should be gambler select? According to the probability theory, the expected gain of the gambler can be calculated as follows.

\[ E[\text{option 1}] = \text{Rs. 50} \]
\[ E[\text{option 2}] = (0.5)(\text{Rs. 100}) \]
So, probabilistically speaking, all the three options have the same expected profit for the gambler, and hence it should not matter which of the three options is chosen by the gambler. However, it has been observed in practice that gamblers do not select option 3 and they prefer option 1 to option 2. Why does it happen? This happens because people are usually averse to risk. In other words, people tend to avoid risk when there is a possibility of losing something that they have. As a consequence, nobody is likely to select option 3, being risky. On the other hand, most people are open to taking risks when there is only the possibility of gaining something. This is the reason, why most gamblers would prefer option 2 to option 1.

Daniel Bernoulli’s Argument

Daniel Bernoulli (1738) recognized the problem with the intuitive notion of ‘expectation of profit’, where one would try to maximize the expected profit. For this consider n possibilities, and assign probabilities $p_i$ ($i = 1, 2, ..., n$) to them along with numbers $M_i$ representing the ‘profit’ that would be obtained if the $i^{th}$ possibility turns out to be true then the expected of profit is given by

$$E(M) = \sum_{i=1}^{n} p_i M_i$$

It is obvious that a human being who acts only for self interest is likely to behave in a way that maximizes the probable profit. It has lead us to a number of paradoxes creating Bernoulli be familiar with that the expectation and desire of more profit cannot always be a reasonable criterion of action.

For instance, consider the situation where the available information leads you to allocate probability 0.51 to the heads in a somewhat biased and unfair coin. You are assigned two actions to choose from. These are (1) to bet all the money you have on the heads for the subsequently toss of the coin, and (2) not to bet at all.

Let $M_0$ be the amount you have now, then your expectation of profit is

$$E(M) = 0.51 M_0 + 0.49(-M_0)$$

$$= 0.02M_0 > 0$$
According to the criterion of expectation of profit, you are expected to gamble when this choice is offered. However, it appeared to be obvious to Bernoulli that nobody who is in the true intelligence would decide the first alternative. This indicates that human general sense regularly rejects the criterion of maximizing probable profit.

Bernoulli’s argument in resolving this paradox was that the true value of a certain amount of money is not simply measured in terms of the amount received. It also depends on how much amount he already has. In other words, Bernoulli argued that it is necessary to realize that the, mathematical expectation of profit is not the similar as its moral expectation. The same idea is expressed and proposed when a modern economist speaks of diminishing marginal utility of money.

A more appealing example is given by St. Petersburg’s paradox, which can be stated as follows. The game consists of a sequence of tosses of a fair coin. Ever toss has a probability of 0.5 to come up to Head and on Tail. The game ends the first time the coin comes up Head. If it takes n tosses for this to happen, you get Rs. $2^n$. The situation is said to be a ‘paradox’ because most persons are agreeable to pay merely a finite, or rather a small, amount in spite of the fact that the expected profit, which is given as follows:

$$\sum_{n=1}^{\infty} (2^{-n})(2^n) = \sum_{n=1}^{\infty} 1 = \infty$$

is unbounded.

Bernoulli also proposed that the “moral value, also called utility in modern economics terminology, of an amount M be supposed to be taken proportional to log M. As a matter fact, Laplace has arrived at the following result in a discussion of the St. Petersburg problem, but without indicating how he arrived at this conclusion”. A person whose total fortune is 200 francs should not reasonably stake more than francs on the play of this game. Now it is not too difficult to verify this result.

For a person who has an amount ‘m’ of money, the cost f(m) may be calculated by equating his present utility with the utility expected if he pays the cost and plays the game as well. This implies that the function f(m) should satisfy:

$$\log m = \sum_{n=1}^{\infty} \frac{1}{2^n} \log(m - f + 2^n)$$
This evaluation on a computer gives \( f(200) = 8.7204 \), showing that Laplace got the result so close without a computer. Similarly, it is interesting to note the following:

\[
\begin{align*}
    f(103) &= 10.9433 \\
    f(104) &= 14.2464 \\
    f(106) &= 20.8754
\end{align*}
\]

These results indicate that still a millionaire should not risk additional to Rs. 21 on the dubious game under any circumstances. Even though such a mathematical result is completely sensible, the logarithmic nature of the utility function should not be taken carelessly and plainly in the cases of enormous little amounts, as has been pointed out by Laplace, and extremely large amounts, as indicated by the following example of Savage (1954):

Suppose you presently own the amount of Rs.10 lakhs and your use of money is proportional to the logarithm of the amount, you then require to be equally inclined to accept and not accept a bet in which, you will be left with Rs. 1000 with probability 0.5 and you will be awarded Rs. 1,000,000,000 with probability 0.5. This kind of a bet would be considered to be disadvantageous and unprofitable for a person with that initial fortune with him. This illustrates that the ‘utility’ for money should increase at a rate that is even slower than the logarithmic function for extremely large amounts.

The essence of Bernoulli’s argument with respect to the gambler’s problem of making the decision under uncertainty is that the decision should be taken in order to maximize the expected value of some portion of the profit, and not the profit itself. Bernoulli called such a function the ‘moral value’. The optimist may describe this as ‘maximizing expected utility’ while the pessimist may describe it as ‘minimizing expected losses, where the loss functions as the negative of the utility function.

1.10 The Logic of Insurance

Consider one more example of a game, which is in some ways similar to St. Petersburg’s game. It is an example of insurance. An oversimplified version of the example is presented here makes some valid and important points. The first argument is that insurance premiums must be sufficiently high for the insurance company to have a guaranteed positive expectation of the profit under all the conditions, favourable as well as unfavourable, covered by the insurance
policy. The second argument is that every rupee the company earns is a rupee a customer loses. The question then remains is: why should anyone ever want to buy insurance?

The core of the argument here is that the particular customer has a utility function for amount of money that can cover a range of values of a magnitude of thousand rupees, but the company is so big that its utility for amount of money is linear in excess of ranges of millions of rupees, suppose ‘P’ be the amount of premium for some suggested insurance policy; let there be n contingencies covered by the insurance policy and let wi be the probability of the i-th contingency with a cost of an amount Li to the company as and when it happens; and let a prospective customer have an initial amount M and have Bernoulli’s logarithmic utility for money. What M stands for is the ‘net worth’ of the customer, and not only the amount of cash be possesses. Then the expected utility for both i.e. the company and the customer, whether he buys or not buys the policy, is shown below:

<table>
<thead>
<tr>
<th>BUY</th>
<th>DON’T BUY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company</td>
<td>( P - \sum Li wi )</td>
</tr>
<tr>
<td>Customer</td>
<td>( \log(M-p) )</td>
</tr>
</tbody>
</table>

As a consequence of the above table, the company is interested offering the policy as long as \( E(L) < P \), whereas the customer is interested in purchasing the policy if

\[
E[\log(M-L)] < \log(M - P)
\]

If the installment is in the range of

\[
E[L] < P < M - \exp\{E[\log(M-L)]\}
\]

it is profitable to both the company and the customer to enter the contract.

The nature of the results discussed above is elementary and important. It is natural for us to think that these results are so good and important to be applied and explored with special focus on Bernoulli and Laplace and there thinking. However, what has actually happened is very different. The ‘frequentist’ way of thinking either ignored this thought or criticized it thus treating it a metaphysical nonsense.

Neyman and Pearson elaborated the form and nature of decision rules as an addition to testing of hypothesis. It became popular for a short period of time, but has now become obsolete because it does not have two fundamental features that are considered to be essential for the problem. Abraham Wald (1950) developed a formulation at a more fundamental level and hence
appears likely to be valid permanently and gives some kind of fundamental justification to Bernoulli’s intuitive ideas. However, what is worth noting is the fact that these hard work were not appreciated universally.

Fisher then shoots his criticism at Neyman and Wald thereby making him as subjects of his criticism. Kendall regarded decision making theory no more than a defect of the American nature as different from the British nature. What is worth noting here is the fact that neither Neyman nor Wald were born, brought up and educated in the U.S.A. History tells us that both of them fled to the U.S.A. from Europe. Fisher regarded decision theory as an abnormal state of minds which is not well versed in a natural science while the fact is that these techniques, procedures and theories were a result of Daniel Bernoulli and Laplace and their character as natural scientists. It is certainly comparable to Fisher’s. According to Kendall approach, Wald gives an idea that inference is simply a particular case of decision making. What is important to keep in mind in this regard is the fact that a clear distinction between inference and decision is maintained in the original formulation of Bernoulli and Laplace. At the same time, while it is necessary to perceive a distinction between inference and decision, it is also important to note that any inference which is not followed by decision is idle, and no scientist would want to undertake the efforts required for conducting inference unless it serves some purpose. The foregoing discussion gives an idea of the nature of challenges faced by Bernoulli and Laplace. However it is necessary to consider Wald’s decision theory.

1.11 Wald’s Decision Theory

Wald’s formulation of decision theory did not appear to have any connection with did not appear to have any connection with probability theory, at least in its initial stages. The formulation begins with enumeration of a set of possible states of nature \{\theta_1, \theta_2, \ldots, \theta_N\} of a finite size. It is theoretically possible to think this set to be countable or even continuous as a useful limiting approximation. For example, in the situation of a quality control problem involving batch inspection, the ‘state of nature’ would be an unknown number of defectives in a group.

The above formulation allows for certain misgivings and misinterpretations to grow and develop right from this elementary stage. What is meant by enumerating all possible ‘states of nature’ is not to describe any real, and therefore apparent and verifiable, property of nature? As a matter of fact, at any given point of time, nature can be in one and only one state. The enumeration is a way of expressing the state of information and concerning the range and
opportunity of possibilities. If the two individuals are asked to do so, they may enumerate the \( \theta_j \) differently and still there need not be any error inconsistency, or contradiction. The objective is only to do the best possible with the available information and it is therefore expected that the person who has better information is bound to make better decisions and that person deserves it too. This apparently arbitrary nature of set of all expected ‘states of nature’ is not, in any way a paradox.

The next step in the formulation of decision theory is to give an illustration of possible decisions \( \{D_1, D_2, \ldots D_k\} \) which are likely to be made. In the problem of quality control with batch inspection, for example, if items in a batch are inspected one after another then at every stage there are three possible decisions as follows:

\[D_1 \equiv \text{Accepting the batch}\]
\[D_2 \equiv \text{Rejecting the batch}\]
\[D_3 \equiv \text{Inspecting one more item}\]

This theory is of no use if we do not act thinking that the decision made is correct. Therefore the enumeration of the possible decisions is a way of expressing the knowledge about what actions are feasible. It is a waste of time as well as of efforts and other resources to consider any decision that we already know to relate to an improbable and impossible course of action.

A particular decision is likely to be eliminated from enumeration on account of one reason or the other. For instance, even though a particular decision, say \( D_1 \), is possible and even easy to carry out, it may be known in advance that it would lead to consequences that are not tolerable. For instance, a car driver take a sudden and sharp curve at any point of time, but his common intelligence normally does not permit him to do that. At this stage, two more points emerge and they are as follows:

1. There is gradation among consequences of a decision. The results of an action can be serious but tolerable.
2. The results of an action, which now represent a decision, also depend on the true state of nature.

Talking about a car driver making a sharp turn, for example, it is also well known that sudden sharp turn does not essentially lead to a disaster. As a matter of fact, sometimes a sudden sharp turn is actually necessary to avoid disaster.
The consideration of possible consequences as a joint effect of the combination of the action (or decision) and the true state of nature, suggest the need of the third concept which is known as the loss function \( L(D_i, \theta_j) \). The loss function represents our judgment regarding the ‘loss’ which can be result of making a decision \( D_i \) if \( \theta_j \) happens to be a true state of nature. If both \( D_i \) and \( \theta_j \) are finite in number, then the loss function is commonly expressed in the form of a matrix with elements denoted by \( L_{ij} = L(D_i, \theta_j) \). A vast amount of literature can be found that deals with criteria for making decisions with only these three elements, namely, a set of states of nature, a set of decisions, and a matrix of the losses for the combination of state of nature and a decision.

The minimax criterion first finds the maximum possible loss \( M_i = \max_j (L_{ij}) \) for every decision \( D_i \) and then selects that \( D_i \) which corresponds to the smallest value of the \( M_i \). In other words, the minimax criterion minimizes the maximum possible loss. This strategy of selecting decision would be reasonable if we assume that nature is so an intelligent adversary that it can anticipate our possible decision and deliberately selects the state of nature which would incur the maximum loss. In the game theory, at least for some games, this is not unrealistic at all and hence minimax decisions are fundamentally important in game theory. However, in decision theory we do not find any intelligent adversary and therefore the minimax criterion reflects the attitude of a pessimist who focuses his attention on the worst possible consequence of his decision and perhaps misses out on the favorable opportunities that the true state of nature may be offering. Never the less, it is also equally unreasonable to go to the other extreme strategy of the ‘minimin’ criterion, where the decision that corresponds to the smallest value of \( m_i = \min_j (L_{ij}) \) is selected. This criterion would make a diametrically opposite, but equally unrealistic, assumption that nature is deliberately trying to help the decision maker by choosing that state which would incur the smallest possible loss.

It should be evident from the above discussion that a reasonable decision criterion should be, to some extent and in some ways, mediates between minimax and minimin, so as to express the belief that nature is neutral towards our actions and their consequences. The literature contains many other criteria suggested by different researchers. Some significant examples are maximin utility (Wald), \( \alpha \)-optimism–pessimism (Hurwicz) and minixam regret (Savage). The standard procedure then has been to analyze any particular criterion that may be proposed from a point of view of satisfying a specified set of qualitative conditions like the two listed next.
1. Transitivity: If $D_1$ is preferable to $D_2$ and $D_2$ is preferable to $D_3$ then $D_1$ must be preferable to $D_3$.

2. Strong Domination: If it known that $L_{ij} < L_{kj}$ for all states of nature $\theta_j$ then decision $D_i$ must always be preferred over $D_k$.

An analysis of this nature can become tedious, even though it is straightforward. It is also understood that there is only one type of decision criteria that fulfils all the conditions. In addition to this, as this type is gained by a very different line of thinking and therefore the considerations such as the ones listed above are not followed any more. A decision theory nevertheless cannot be said to be complete with only the $D_i$, $\theta_j$ and $L_{ij}$. Besides, there isevidence, i.e. $E$, which is known as being related to the decision problem under concern and it is compulsory to be trained to apply the evidence $E$ to the theory. Wald proposed and developed the decision theory up to this particular point and then made a long, difficult and apparently unnecessary, mathematical argument. The decision maker defines, a strategy $S$, which consists of a set of rules relating the form if a decision maker receives new evidence $E_k$, and then he will make decision $D_i$. The decision maker, in principle, first makes a list of all conceivable strategies and then eliminates those that are considered to be undesirable according to the following criteria. Let the equation

\[
p(D_i | \theta_j, S) = \sum_k p(D_i | E_k, \theta_j, S) p(E_k | \theta_j)
\]

denote the probability that the strategy $S$ would lead towards making a decision $D_i$ when $\theta_j$ is the right state of nature. Then the risk projected by $\theta_j$ for the strategy $S$, is defined as the possible loss with look upon to this probability distribution. That is

\[
R_j(S) = \sum_i p(D_i | \theta_j, S) L_{ij}
\]

In this set-up, a strategy $S$ is measured to be allowable if there is no other strategy $S'$ that satisfies the following form

\[
R_j(S') = R_j(S), \quad \text{for all } j
\]

Furthermore, the strategy $S$ is said to be inadmissible if a strategy $S'$ exists for which the above equation is strict for at least one $\theta_j$. The concepts of risk and admissibility are obviously criteria based on sampling theory, and not on the Bayesian theory, because they require the sampling distribution only. Wald also thought on the lines of the theory known as a sampling
theory, and arrived to the conclusion that the optimal strategy must be searched only in the class of permissible strategies.

The main objective of Wald’s theory is to describe the group of admissible or permissible strategies. Wald desired to do this especially in numerical forms so that every admissible strategy may be established simply by performing a specified definite procedure. The basic theorem dealing with this situation is none other than Wald’s complete class theorem, that brings out the result was shocking to the experts in sampling theory. The phrase ‘complete class’ is defined in a strange way. What Wald wanted to identify was only the set of all admissible and permissible rules. Berger (1985) considers this class a ‘minimal complete class’. Wald considered the problem of proving existence of a ‘complete class’ to be a highly non-trivial mathematical problem. Wald also went ahead to find an algorithm for making rules in this class.

However, there now appears to be unnecessary complication and they only signify an inappropriate definition of the term ‘admissible’. It is known in many situations that any inadmissible decision rule is surprisingly preferable to any admissible decision rule, mainly on account of the fact that the criterion of admissibility totally ignores any previous information. This reasoning show how the terms like ‘admissible’ and ‘unbiased’ can be misleading because they are not neutral in ethical or moral sense. These terms indicate something like a good character and good intention. Now it is known that there is no need to restrict the attention only to the set of admissible rules, because there is a very different line of thinking leading towards the rules which are suitable in the practical humanity.

This leads us to a fundamental question of what makes a decision process difficult. If the state of nature is understood correctly, then there is no problem at all. This is so because, in such a situation, the best decision $D_i$ gives the minimum of $L_{ij}$ where $\theta_j$ is known to be the true state of nature. As a result, what can be said is that once the loss function is specified, the uncertainty about the expected decision arises out of uncertainty about the true state of nature. Therefore, whether the decision minimizing $L_{ij}$ is best or not depends on how confidently we believe that $\theta_j$ is the true state of nature. In other words, a judgment is required regarding how plausible $\theta_j$ is. So the question can be stated as follows. Considering all the available evidence, what is the probability $p_j$ that $\theta_j$ is the true state of nature? This question cannot be asked to a sampling theory expert, who interprets probability as “long-run relative frequency in a random experiment”. If we use this interpretation of probability, it is not appropriate to think of the
probability of $\theta_j$ because the state of nature is not a random variable. In this way, it should be clear by now that if we stick to the sampling theory perspective of probability, we shall arrive at the conclusion that it is not possible to apply the probability theory to the decision problem. It is this kind of approach that has made many statisticians in the earlier stage of the last century to drive problems of assessing and judging parameters and testing of hypotheses to a new area that is known as Statistical Inference. This new area was regarded as dissimilar from probability theory and was considered to be based on entirely different principles.

1.12 Problem on Hand

The main problem addressed in the proposed thesis is related to development of new methods for efficient decision making when the situation is complex, the risk is high and uncertainty is multifold. Complexity of situation demands that the decision can not be taken simplistically. It must be elaborate and to be taken step by step. The high risk requires that a more sensitive and reasonable function may be necessary to define and introduce for identifying or developing better decision making algorithms. One of the proposed methods is based on decision trees. Today, most of the decision tree algorithms are more like classification trees, because they are supervised in nature. They do not take into account the loss or risk function, but only minimize the impurity of each resulting branch. The proposed thesis aims at developing a tree-based decision making process that would satisfy optimality criteria based on a risk function. It is also proposed to test the decision tree for its sensitivity, because a robust decision tree is more useful than a sensitive one.

1.13 Research Objectives

1. To define new loss functions according to the nature of the predictor variable and the response variable.
2. To define a tree-forming method on the basis of a nested decision procedure, where one decision rule is applied after another.
3. To derive the optimal decision rule taking into account the data features relevant to the decision at every node of a resulting decision tree.
4. To measure the efficiency of the resulting decision tree at leaf level and at the level of the entire tree.
5. To identify predictor variables that is significant in building the decision tree. In addition, the significant predictors will be remarked according to their relative importance in the model.
6. To measure and specify the contribution of every significant predictor variable in building the decision tree.
7. To test sensitivity of the resulting decision tree by repeated analysis of permuted datasets.

1.14 Scope

Decision theory is applicable in situations where decisions are to be made under uncertainty. Therefore, the present research has scope in all those situations where decisions are to be made under dynamic uncertainties, rather than static ones. Scope for data-driven and tree-based decision theory is almost in every field of management, because in most situations of decision making the conditions are not universal, but only restricted to a particular case.

1.15 Hypothesis

The hypothesis of the chapter foundations of decision theory are multidisciplinary in nature, and are spread over disciplines of philosophy, psychology, economics, social sciences, management science, computer science and engineering in addition to its roots in statistics. The methodology is mostly analytical. The chapter aims to establish the foundations of decision theory in a coherent manner, so that later developments can be both understood and appreciated in comparison to the foundations.

The second hypothesis in this chapter is that decisions are more sensitive to experience (that is, data) than to the unknown state of nature (that is, the parameter). The methodology is to find ways to make decisions depend heavily on data by developing supervised decision systems.

The hypothesis of the chapter Objectivity and Subjectivity in Decision Making is that pure or total objectivity is impossible in any decision making process. This is so because decisions are taken only on the basis of available information, which cannot be perfect. The methodology of the chapter is to develop methods to identify subjectivity in decision making, so that the effect of subjectivity can be eliminated, or minimized, from the recommended decision rule.

The hypothesis of the chapter Sequential Decision Making is that most decision problems are not simple enough to be solved in a single step. Further, they cannot be solved by solving their parts arbitrarily. It is therefore necessary to solve them sequentially. The methodology is
break up the problem and establish the dependency among parts of the problem in order to identify the order in the sequential approach to solving the problem.

The hypothesis of the chapter, Construction of Decision Trees, using conditional Decision Rules is that every complex decision rule can be presented in the form of a decision tree. The branches represent different possible actions and the end node of edges indicates the consequences of those actions at every node. The depth of the tree is determined by the number of sub-problems, where every sub-problem resulting in branching. The methodology involves analyzing the sub problem to check the presence of any hierarchy of order, so that the appropriate tree can be constructed. Most of sequential decision problems are naturally represented by a tree, but every tree need not represent a sequential decision procedure.

The hypothesis of the chapter ‘Optimal Decision Tree Algorithms’ is that decision trees can be optimized by specifying optimality criteria. It has to be decided whether the resulting decision tree should be globally optimal or if it is good enough to have it only locally optimal. An algorithm for global optimality will obviously have a higher level of complexity. The methodology is typically either iterative or recursive. It is necessary to involve dynamic programming if global optimality is required.