CHAPTER 1

INTRODUCTION

1.1 STOCHASTIC PROCESS

In the past periods, it is realized that the probabilistic models are as realistic as deterministic models for the design and development of Stochastic processes. The stochastic processes are mainly used in engineering sciences, biological sciences, socio and economic systems. To study these systems it is important to examine the probability laws governing the behaviour of these systems.

Mathematical models can be divided into two categories, namely deterministic and stochastic models. Deterministic models are usually expressed in terms of mathematical equations with initial and boundary conditions which exactly predict the development of a system. Stochastic models are given by random variables whose outcomes are uncertain where we can compute only the probabilities of possible outcomes. Though useful, the division line between deterministic and stochastic modeling approaches is far from being sharp.
In the study of any system, the main work is to derive mathematical models for the responses in terms of input variables as well as time. Almost all these models are Stochastic models, since most of the variables are subjected to random variations and most of the measurements of the responses are subjected to random measurement errors.

Stochastic models are used in several fields of research. Some of the models like traffic flow models, queuing models, reliability models, and spatial-temporal models are used in engineering sciences. In computer science, the queuing theory is used to compare the performance of different computer systems. Stochastic models are also applied in the medical systems such as AIDS, cancer and genetics as well as for Social Economic systems.

1.1.1 DEFINITION OF STOCHASTIC PROCESSES

A family of random variables \{X(t), t \in T\} is called stochastic processes for each \( t \in T \), where \( T \) is the index set of the process, \( X(t) \) is a random variable. An element of \( T \) is usually referred as a time parameter. The state space of the process is the set of all possible values that random variables \( X(t) \) can assume. Each of these values is called the state of the process.

A Markov chain model developed by the Russian mathematician Andre A. Markov in 1905 are a particular class of probabilistic model.
and is known as Stochastic processes, in which the current state of the system depends on all of its previous states. But in Markov process, the current states of the process depend only on immediate preceding state. The Markov Process is a system that can be in one of several numbered states, and pass from one state to another for each time step according to fixed probabilities.

1.1.2 PARAMETER SPACE AND STATE SPACE

The random variable which is given with respect to the parameter \( t \), is usually denoted by \( \{ X_t \} \). The parameter \( t \) is usually referred to time. A set of all possible values of the indexing parameter is called the parameter space. For any fixed \( t \), \( X_t \) is a random variable. The possible values of this random variable \( X_t \) are called the states of the process. That is, the state of the process is defined as the position of the process at any time, i.e., when \( X_t = \omega \), the process is in a set \( \omega \) at a time \( t \). The set of all possible values, i.e., the set of all states is called the state space of the process and is denoted by \( S \). Based on the index set \( T \) and the state space \( S \), a Stochastic Process is usually specified in the form \( \{ X_t : t \in T \} \) over \( S \).

1.2 MARKOV CHAIN

The Stochastic process \( X = \{X_n : n \in \mathcal{N}\} \) is called a Markov Chain (MC) provided that \( P[X_{n+1} = j | X_0, X_1, ..., X_n] = P[X_{n+1} = j | X_n] \) for all \( j \in S \).
where \( S \) is the countable state space and \( n \in \mathbb{N} \). A Markov chain is a sequence of random variables such that for any \( n \), the next state \( X_{n+1} \), of the process is independent of the past states \( X_0, X_1, \ldots, X_{n-1} \), provided that the present state \( X_n \) be known.

A Markov chain \( X \) is said to be time-homogenous if the conditional probability \( P[X_{n+1} = j \mid X_n = i] = P_{ij} \), \( i, j \in S \) is independent of \( n \) and \( S \) is the countable state space. The probabilities \( P_{ij} \) are then called the transition probabilities for the Markov chain \( X \).

It is customary to arrange the \( P_{ij} \) or \( P(i,j) = P_{ij} \) into a square array and to be called the resulting matrix \( P = (P_{ij}) \) the transition matrix of the Markov chain \( X \); for any \( i, j \in S \), \( P_{ij} \geq 0 \), and for each \( i \in \mathcal{E}, \sum_{j \in \mathcal{E}} P_{ij} = 1 \) for any \( m \in \mathbb{N} \). \( P[X_{n+m} = j \mid X_n = i] = P_{ij}^m \); \( m \)th power of \( P_{ij} \).

### 1.2.1 CLASSIFICATION OF MARKOV PROCESSES

<table>
<thead>
<tr>
<th>Time</th>
<th>State Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discrete</td>
<td>Discrete time Markov Chain</td>
</tr>
<tr>
<td></td>
<td>Continuous time Markov Process</td>
</tr>
<tr>
<td>Continuous</td>
<td>Continuous time Markov Chain</td>
</tr>
<tr>
<td></td>
<td>Continuous time Markov Process</td>
</tr>
</tbody>
</table>
For a discrete time Markov chain, the process is assumed as making state transitions at times $t_n$, $n = 1, 2, 3, \ldots$. Thus, the discrete time Markov chain $\{X_n\}$, $(X_n$ for $X(t_n))$ starts with an initial state, say $i$, when $t = t_0 (X_0 = i)$ and makes a state transition at the next step (time in the sequence); that is, when $t = t_1$, so that $X_1 = j$, etc. The one-step transition probabilities is defined by $P[X_{n+1} = j | X_n = i]$, for every $n$ and $i, j = 0, 1, 2, \ldots$.

### 1.3 CONCEPT OF FUZZY LOGIC

Fuzzy logic refers to a logical system that generalizes the classical two value logic for reasoning under uncertainty. It is a system of computing and approximate reasoning based on collection of theories and technologies that employ fuzzy sets, which are classes of objects without sharp boundaries where membership is a matter of degree [Yen and Langari, 1999]. The mathematical background of fuzzy logic is associated to the fuzzy set theory, which is the extension of the classical set theory.

The basic concept underlying fuzzy logic is that of a linguistic variable, whose values are words rather than numbers. Although words are often less precise than numbers and their use is closer to human perception. Another basic concept in fuzzy logic is “if-then” rule called a fuzzy rule or fuzzy inference system which plays an
important role in most of its applications and it is based on natural language used by people on a daily basis.

Fuzzy logic provides a simple way to arrive at definite conclusion based upon vague, ambiguous, imprecise and noisy, or missing information. Fuzzy logic technology has achieved the impressive success in diverse engineering applications ranging from mass market consumer products to sophisticated decision and control problems.

1.3.1 EXAMPLE OF FUZZY SET

A fuzzy set is a set containing elements that have varying degrees of membership in the set. Fuzzy set provides a suitable point of departure for the construction of conceptual framework which parallels in many respects of the framework used in the case of ordinary sets, but is more general than the latter and potentially, may prove to have a much wider scope of applicability, particularly in the field of pattern classification and information processing. A set is a collection of some objects of the universe. To fix the elements of the set A, naturally a condition is given. Objects satisfying the condition belong to the set A and objects not satisfying the condition do not belong to A.

When a set A is stated and if we consider any object in the universe, then there are only two possibilities. The object chosen satisfies the condition and hence it is an element of A, or the object
chosen does not satisfy the condition and hence it is not an element of A. So the entire set theoretical concept is based on the assumption that a given condition divides the entire universe exactly into two parts, one part contains elements satisfying the condition and another part contains elements not satisfying the condition. So, if we consider any object and if we ask the question whether the condition is satisfied by the object or not then the answer must be either YES or NO. This is called the Binary logic.

But in general there are some attributes, the existence of which in an individual can not be determined by yes or no. As an example, let A be the collection of all talented students in a class having 100 students. Suppose a particular student always scores 40% in any exam then is he talented or not? We don't want to say that he is talented, at the same time we don't want to close the matter by saying that he is not talented. Suppose a student consistently scores 40% marks, is he talented or not? We can't say. So the two way logic (yes or no logic) (True or False logic) doesn't work. The drawback of conventional sets is that many concepts encountered in the real world cannot always be described exclusively by their membership and non-membership in sets.

What best can be said about this student, we can say he is 40% talented. So he belongs to the set of all talented students with a measure of say 40%. Like that all the hundred students can be given
the percentage. So consider a universe which contains all the objects of consideration, given an attribute, each member in the considered universe is given a number (percentage) which measures the existence of that attribute in it.

The first publication on fuzzy set theory by [Zadeh, 1965] shows the intention to generalise the classical notion of a set and a statement to accommodate fuzziness. The word “imprecision” here is meant in the sense of vagueness rather than the lack of knowledge about the value of a parameter. Fuzzy set theory provides a strict mathematical framework in which vague conceptual phenomena can be precisely and rigorously studied [Zimmermann, 1991]. The notion of a fuzzy set is completely non-statistical in nature [Zadeh, 1965]. Fuzzy set theory takes the same logical approach as researchers use the concept of classical set theory. In the classical set theory, as soon as the two-valued characteristic function has been defined and adopted, rigorous mathematics follows. However, in the fuzzy set case, as soon as a multi-valued membership function has been chosen, a rigorous mathematical theory can be developed.

Let $X$ be the considered universe. A be an attribute for any $x$ in $X$ we have a measure say $\mu(x)$. Hence any attribute gives a map $\mu : X \rightarrow \mathbb{R}$ where $\mu(x)$ measures the existence of the attribute in $x$. This is defined as a fuzzy set.
A is fuzzy set in \( X \) means given any \( x \) in \( X \) we cannot say \( x \) belongs to \( A \), or \( x \) does not belong to \( A \), but we say that \( x \) belongs to \( A \) with a measure of \( \mu(x) \), \( \mu \) is called the membership function. We assume \( \mu(x) \) lies in \([0, 1]\) for all \( x \). In general a fuzzy set on \( X \) is nothing but a map (function) \( X \to [0, 1] \).

### 1.3.2 UNIVERSE OF DISCOURSE

All elements in a set are taken from a *universe of discourse* or *universe set* that contains all the elements that can be taken into consideration when the set is formed. In reality there is no such thing as a set or a fuzzy set because all sets are subsets of some universe set, even though the term ‘set’ is predominantly used. In the fuzzy case, each element in the universe set is a member of the fuzzy set to some degree, even zero. The set of elements that have a non-zero membership is referred to as the support. We will use the notation \( U \) for the universe set.

### 1.3.3 LINGUISTIC HEDGES

Humans communicate with their own natural language by referring to previous mental images with rather vague but simple terms. Therefore any attempt to model the human thought process as expressed in our communications with one another must be preceded by models that attempt to emulate our natural language. Natural language consists of fundamental terms called "atomic terms". Examples of some atomic terms are “medium”, “young” and
“beautiful”, etc. Collection of atomic terms are called composite terms. Examples of composite terms are “Very slow car”, “Slightly Young student”, “fairly beautiful lady”, etc. Suppose we define the atomic terms and sets of atomic terms to exist as elements and sets on a universe of natural language terms, say universe X. Further more let us define another universe, called Y, as a universe of cognitive interpretations, or meanings. Though it may seem easy to envision the universe of terms, it may be difficult to ponder a universe of “interpretations”. The atomic terms are called linguistic variable in Fuzzy set theory.

In linguistics, fundamental atomic terms are often modified with noun or verbs like very, low, slightly, more-or-less, fairly, almost, barely, mostly, roughly, approximately, etc. These modifiers are called “linguistic hedges,” that is, the singular meaning of an atomic term is modified or hedged from its original interpretation. When we use fuzzy set for interpretation, the linguistic hedges have the effect of modifying the membership function for a basic atomic term. The following are examples of hedges used in different types of variables:

- very, quite, extremely - general modifiers
- quite true, mostly true - truth value modifiers
- likely, not very likely - probability modifiers
- low, medium, high - quantity modifiers
For example, the temperature in fuzzy linguistic variable with hedges can be characterized as a fuzzy set as shown in figure 1.1.

![Figure 1.1 Representation of fuzzy linguistic variable](image)

We say today it is hot or today the temperature is high. In fact, we talk in terms of vagueness/ambiguity. Figure 1.1 shows that the variations of low, medium and high temperature in degree centigrade with membership function. The adjective low temperature refers to 0-20 degree centigrade; medium temperature refers to 10 – 40 degree centigrade and high refers to above 30 degree centigrade. There is an overlap between all these three zones. These are defined as regards to membership function.

### 1.3.4 OPERATIONS IN FUZZY SETS

In crisp set theory we have the concepts of union and intersection of two sets. This concept should be extended to fuzzy sets.
Also the concept of complement of a set in crisp set theory should be extended to fuzzy sets [Sivanandam et al., 2007].

**Definition:** Let $A$ and $B$ be two fuzzy sets on a nonempty set $X$. The union of $A$ and $B$ denote as $A \cup B$ is defined as

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)), \quad \forall x \in X,$$

where $\mu_{A \cup B}$ is the membership function of $A \cup B$ is map from $X$ to $[0, 1]$. Hence $A \cup B$ is a fuzzy set on $X$. This is called the **standard union of two fuzzy sets**.

![Figure 1.2: Union of Fuzzy sets $A$ and $B$.](image)

**Definition:** Let $A$ and $B$ be two fuzzy sets on a nonempty set $X$. The intersection of $A$ and $B$ denoted as $A \cap B$ is defined by

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)), \quad \forall x \in X$$

where $\mu_{A \cap B}$ is the membership function of $A \cap B$ is a map from $X$ to $[0, 1]$. Hence $A \cap B$ is a fuzzy set on $X$. This is called the **standard intersection of two fuzzy sets**.
Definition: Let $A$ be a fuzzy set on nonempty set $X$. The complement of $A$ denoted as $A^c$ is defined as $\mu_{A^c} = 1 - \mu_A(x), \forall x \in X$. Clearly $A^c$ is a map from $X$ to $[0, 1]$. Hence $A^c$ is a fuzzy set. This is called the *standard complement*.

Definition: Let $A$ and $B$ be two fuzzy sets on a non empty set $X$. The difference denoted by $A - B$ is defined as $A - B = A \cap B^c$

1.3.5 **MEMBERSHIP FUNCTIONS**

A membership function is a curve that defines how each point in the input space is mapped to a membership value between 0 and 1. Sometimes the input space refers to universe of discourse.
**Definition:** For any set $X$, a membership function $\mu_A$ on $X$ is any function from $X$ to the real unit interval $[0, 1]$. Membership functions on $X$ represent fuzzy subsets of $X$. The membership function which represents a fuzzy set $\tilde{A}$ is usually denoted by $\mu_A$. For an element $x$ of $X$, the value $\mu_A(x)$ is called the *membership degree* of $x$ in the fuzzy set $\tilde{A}$. The membership degree $\mu_A(x)$ quantifies the grade of membership of the element $x$ to the fuzzy set $\tilde{A}$. The value 0 means that $x$ is not a member of the fuzzy set; the value 1 means that $x$ is fully a member of the fuzzy set. The values between 0 and 1 characterize fuzzy members, which belong to the fuzzy set partially.

![Figure 1.5: Membership function of a fuzzy set](image)

The general membership function of a fuzzy set can be defined as the functions which take values in an arbitrary fixed algebra or structure $L$; usually it is required that $L$ be at least a poset or lattice. The usual membership functions with values in $[0, 1]$ are then called $[0, 1]$-valued membership functions.
Mathematically, the above membership function can be defined as,

$$\mu_A(x) = \begin{cases} 
0, & x < a \\
\frac{x-a}{b-a}, & a \leq x \leq b \\
1, & x > b 
\end{cases}$$

In the case where a fuzzy set $A$ is a conventional (crisp) set, the corresponding membership function can be reduced to

$$\mu_A(x) = \begin{cases} 
1, & x \in A \\
0, & x \notin A 
\end{cases}$$

The above function has only two outputs, 0 or 1. Whenever $\mu_A(x) = 1$, $x$ is a member of $A$, if $\mu_A(x) = 0$, $x$ is declared a non-member of $A$.

**Triangular Fuzzy Number:** A triangular fuzzy number $A$ is a fuzzy number with a piecewise linear membership function $\mu_A(x)$ and is defined by

$$\mu_A(x) = \begin{cases} 
\frac{x-a_1}{a_2-a_1}, & a_1 \leq x \leq a_2 \\
\frac{a_3-x_1}{a_3-a_2}, & a_2 \leq x \leq a_3 \\
0, & otherwise 
\end{cases}$$
Figure 1.6: Triangular Membership function of a fuzzy set

Triangular fuzzy number are more often used when fuzziness exist on both sides of a single value/parameter/factor.

**Trapezoidal Fuzzy Number:** A trapezoidal fuzzy number $A$ is a fuzzy number with a piecewise linear membership function $\mu_A$ and is defined by

$$
\mu_A(x) = \begin{cases} 
0, & \text{when } x < a_1 \text{ and } x > a_4 \\
\frac{x-a_1}{a_2-a_1}, & a_1 \leq x \leq a_2 \\
1, & a_2 \leq x \leq a_3 \\
\frac{a_4-x}{a_4-a_3}, & a_3 \leq x \leq a_4 
\end{cases}
$$
Trapezoidal fuzzy numbers are generally used when fuzziness exists on both side of an interval.

**Gaussian Fuzzy Number:** A Gaussian membership function is defined by

\[ G(x : \mu, \sigma) = \exp\left[-\left(\frac{x - \mu}{\sqrt{2}\sigma}\right)^2\right] \]

where the parameters \( \mu \) and \( \sigma \) control the center and width of the membership function.

### 1.3.6 Defuzzification Method

Defuzzification is the conversion of a fuzzy quantity to precise quantity, just as fuzzification is the conversion of a precise quantity to a fuzzy quantity. In practice, the output of the defuzzifier process is a
single value from the set. There are several built-in defuzzifier methods. The centre of gravity method is the most commonly used for extracting a crisp value from a fuzzy set. This method calculates the weighted average of the elements in the support set. The bisector method focuses on the axis of the vertical line which divides the area under the diagram into two equal parts. The mean of maxima method chooses the point by taking the mean of the maximal memberships. The smallest maximum and largest maximum methods choose either the lower or upper boundary of the maximal membership. For computational complexity, the gravity bisector methods are categorized.

1.4 TIME SERIES MODELING

A basic assumption in any time series modeling is that some aspects of the past pattern will continue to remain in the future. Under this set up, the time series process is based on past values of the main variable but not on explanatory variables which may affect the system. Hence it is assumed that information about the past is available in the form of numerical data.

Time series analysis and its applications have become increasingly important in various fields of research, such as business, economics, engineering, medicine, social sciences and politics. This analysis can be used to carry out different goals such as descriptive analysis, spectral analysis, forecasting, intervention analysis and
19

explanative analysis. Since [Box and Jenkins, 1970] published the book entitled Time Series Analysis: Forecasting and Control, a number of books and a huge number of research papers have been published in this area. In the classical theory of time series analysis, one used to assume that the structure of the series can be represented by linear time series models, for example, the autoregressive model (AR), moving average model (MA) and autoregressive moving average model (ARMA), and autoregressive integrated moving average model (ARIMA) by taking into account the seasonality effect.

1.5   **FUZZY TIME SERIES**

Fuzzy time series was proposed by [Song and Chissom, 1993]. It defines that imprecise data at equally spaced discrete time points which are modeled as fuzzy variables. The set of this discrete fuzzy data forms a fuzzy time series and also it defined that chronological sequences of imprecise data are considered as time series with fuzzy data. A time series with fuzzy data is referred to as fuzzy time series.

Several Fuzzy Time Series (FTS) models have been investigated in scientific literature during the last two decades. Among these, the most accurate Fuzzy Time Series models found in literature are the high order models. However, three fundamental issues need to be resolved with regard to the high order models. First, current prediction methods are not able to provide satisfactory accuracy rates for defuzzified outputs. Second, data becomes under utilization as the
order increases. Third, forecast accuracy is sensitive to selected interval partitions.

Fuzzy time series is quite common in our daily life. For example, one usually use the linguistic terms such as ‘good’, ‘bad’, ‘not very good and so on to express one’s mood or feeling. By recording such observations, one will have a dynamic process whose observations are linguistic or fuzzy sets.

1.6 ORGANISATION OF THE THESIS

This thesis is comprised into seven chapters which gives the reader an extensive knowledge and research view of Stochastic modeling in fuzzy time series. It also describes the concepts and definitions of the Stochastic Processes, Markov Chain, Time series analysis, Fuzzy logic, Fuzzy time series and their relevant applications.

In chapter II consists of review of research papers related to Markov modeling in fuzzy time series. The survey has been done in chronological order to give the reader a detailed review of these years with contemporary papers such as fuzzy set operations, fuzzy logical relations, defuzzification, computational algorithm, identification of outlier in fuzzy time series, Markov model, Hidden Markov model,
Heuristic model, Higher order multivariate Markov chain model, ARIMA Models, Entropy based Fuzzy time series with applications and other related topics.

In chapter III is discussed for number of accidents data set forecasting using Markov model in fuzzy time series. The forecasting accuracy is studied by removing the identified outliers in the data using Cook's distance and Studentized residual test. After removing the identified outlier observation, Markov modeling is adopted for the rest of the data set. A modified fuzzy time series method is based on transition probability vector membership function. The proposed method gets better forecasting results. It is experimentally found that the proposed method minimizes the average forecasting error compared with other known existing methods. The advantage of this method is to giving better accuracy for average forecasting error compared to time-invariant models which is supported by numerical and graphical representation.

In chapter IV attempts are made to improve the accuracy of forecasting Hidden Markov model in fuzzy time series for bivariate data set. The computation is carried out using fuzzy sets and transition probability vectors. Moreover, by applying the Monte Carlo simulation method when estimating the forecasting result, the randomness of the forecasting are made more reasonable to test the
effectiveness of the resulting stochastic model. As documented in the experiment results, the proposed Hidden Markov model in fuzzy time series minimizes the average forecasting error compared with other known existing methods.

In chapter V multivariate Markov chain model is used to represent the behaviour of multiple time series data generated by similar data set or same data set is discussed. The impact of atmospheric CO$_2$ emission upon global surface temperature is analyzed. A time series model is determined by applying Fuzzy logic and an attempt is made to forecast global surface temperature using higher order multivariate Markov Models. Reliability of the randomness of the forecast value is studied by applying simulation technique. Numerical computation has been carried out in support of the theoretical findings and it is justified that the proposed method gives better forecasting results.

In chapter VI described a computational algorithm in fuzzy time series and it is emphasized that the forecast uses only historical data. The computational algorithm applied for accident data during the year 1981-2005 in India. The concentration is to reduce the average forecasting error and the comparison with the existing methods.

Chapter VII describes the summary of the chapters from I to VI.