INTRODUCTION

1. Introduction

Since last century there have been marked changes in the approach to scientific enquiries. There has been greater realization that probability models are more realistic than deterministic models in many situations. Many a phenomenon occurring in physical and life sciences are studied now not only as random phenomenon but also as one changing with time or space. Such phenomena can be described in terms of stochastic process.

From the non-mathematical point of view a stochastic process is a probability process i.e., any process running along in time and controlled by a probabilistic law. A family of random variables \( \{ X(t), t \in T \} \), which are functions of time \( t \) (say) is known as a stochastic process. Here \( X(t) \) is in practice the observation at time \( t \) and \( T \) is the time range involved.

Most classical problems in probability involve only finitely many random variables, that is the corresponding range \( T \) of \( t \) is a finite set. If \( T \) is an infinite sequence, the process \( \{ X(t), t \in T \} \)

\[
\begin{align*}
X_m & \quad m=1, 2, \ldots \quad (m \rightarrow \infty) \\
X_0 & \quad X_1, X_2, \ldots
\end{align*}
\]
is known as discrete parameter process. On the other hand if $T$ is an interval then $\{X_t, t \in T\}$ becomes a continuous parameter family and the process is called continuous parameter process. In this case the general sample point is a function of $t$ defined on an interval, whereas in the discrete parameter case the general sample point is a sequence.

In the classical theory of statistics, we study the models in which observations are independent and identically distributed (i.i.d.). In many cases we come across the process where members are not mutually independent. The present value of a random variable may depend on just the preceding value and such a dependence is known as Markov dependence.

Markov process is named after A.A. Markov who introduced the concept in 1906 with discrete parameter and finite number of states. The denumerable case was launched by Kolomogorov in 1937 followed closely by Doeblin whose contribution pervade all parts of Markov theory. The continuous parameter stochastic process is rigorously introduced in Wigner (1923). These processes are now known as Wigner processes. Fundamental work on continuous parameter chains was done by Doob (1942, 1949), Paul Levy (1951) with his unique intuition, drew a comprehensive picture of the area.

A Markov process is defined as a stochastic process $\{X(t), t \in T\}$ such that for the given value of $X(s)$, the distribution function of the values of $X(t)$ ($t > s$) does not depend on the values of $X(u)$, ($u < s$).
The set of all possible values of individual random variable $X_n (X(t))$ is known as the state space.

Markov processes in which time parameter is discrete and the state space is also discrete are known as Markov chains. Markov processes in which time parameter is continuous and the state space is discrete are known as continuous time Markov chains, and those in which both time parameter and the state space are continuous and also if the sample paths are continuous are known as diffusion processes. For the discussion of all these processes we may refer to any introductory book on stochastic processes (e.g. Doob (1953), Medhi (1982) etc among many others).

Adke and Manjunath (1984) have discussed Markov process with continuous time parameter and discrete state space. They have defined such processes as discrete Markov processes. If the state space is finite then they are known as finite Markov processes. Adke and Manjunath (1984) have given alternative definitions for Markov processes and established their equivalence. They have discussed the classification of states for such discrete Markov process.

Let $P_{ij}(t)$ denote a transition probability function i.e.,

$$P_{ij}(t) = \Pr \{ X(t) = j \mid X(0) = i \}$$

(1.1)

$(i, j) \in S$, where $S$ is the state space. Let $\lambda_{ij}$ denote the rate of transition from state $i$ to $j$.  

...
\[
\lambda_{ij} = \lim_{t \to 0} \frac{P_{ij}(t)}{t}, \quad (i \neq j) \quad (1.2)
\]

and
\[
\lambda_i = \lim_{t \to 0} \frac{1 - P_{ii}(t)}{t} = \sum_{j=1}^{m} \lambda_{ij} > 0. \quad (1.3)
\]

Here \( \lambda_i \) is the rate of transition from state \( i \) to other possible states.

A state \( i \) for which \( 0 < \lambda_i < \infty \) is called stable. It is absorbing if \( \lambda_i = 0 \). If then the waiting time in state \( i \) is a random variable whose distribution function is exponential.

A state \( i \) for which \( \lambda_i = \infty \) is called an instantaneous state and the expected waiting time in such a state is zero.

Now, consider a stochastic process known as Markov renewal process in which change of state occurs according to a Markov chain and the time interval between consecutive transitions is a random variable whose distribution may depend on the state from which transition takes place as well as on the state to which transition takes place. Such a process becomes a Markov process if the distribution of the sojourn times are all exponential, independent of the next state; it becomes a Markov chain if the sojourn times are equal to one and it becomes a renewal process if there is only one state.

Suppose we have for each \( n \in \mathbb{N} \), a random variable \( X_n \), taking values in a countable set \( E \) and a random variable \( t_n \) taking values in \( \mathbb{R}^+ = [0, +\infty] \) such that \( 0 = t_0 < t_1 < t_2 \leq \ldots \). The stochastic process \( (X, T) = \{X_n, t_n; n \in \mathbb{N}\} \) is said to be a Markov renewal
process with the state space $\mathcal{S}$, provided that
\[
P \left[ X_{n+1} = J \ , \ t_{n+1} - t_n \leq \bigcup_{i \neq J} X_n \ , \ X_1 , \ldots , X_n , t_0 , t_1 , \ldots , t_n \right] \\
= P \left[ X_{n+1} = J \ , \ t_{n+1} - t_n \leq \bigcup_{i \neq J} X_n \right]
\] (1.4)
for all $i, J \in S$ and $t \in \mathbb{R}^+$. If $(X, T)$ is a time homogeneous Markov renewal process then for any $i, J \in S$, $t \in \mathbb{R}^+$.
\[
P \left[ X_{n+1} = J \ , \ t_{n+1} - t_n \leq \bigcup_{i \neq J} X_n \right] = Q_{ij}(t)
\] (1.5)
is independent of $n$. The family of probabilities $Q = \{ Q_{ij}(t) \ ; \ i, J \in S \ , \ t \in \mathbb{R}^+ \}$ is called semi-Markov kernel over $S$.

Thus we have a Markov chain $X = \{X_n \ ; \ n \in \mathbb{N} \}$ with transition probabilities
\[
P_{ij} = \lim_{t \to \infty} Q_{ij}(t)
\] (1.6)
associated with Markov renewal process $(X, T)$. Evidently $P_{ij} \geq 0$ and $\sum_{j \in S} P_{ij} = 1$. The increments $(t_1 - t_0), (t_2 - t_1), \ldots$ are all conditionally independent given the Markov chain $X_0, X_1, \ldots$ with the distribution of $(t_{n+1} - t_n)$ depending only on $X_n$ and $X_{n+1}$ (cf. Cinlar (1975)).

Another convenient description of a Markov renewal process is provided by the process $Y = \{Y(t) \ ; \ t > 0 \}$ defined by putting for each $t \geq 0$
\[
Y(t) = X_n \ \text{on} \ t_n \leq t < t_{n+1}
\] (1.7)
The stochastic process $Y = \{ Y(t), \ t > 0 \}$ defined by (1.7) is called minimal semi-Markov process associated with $(X, T)$. Thus a semi-Markov process $\{ Y(t), \ t > 0 \}$ is a stochastic process with continuous time parameter and discrete state space, in which change of state occurs in accordance with Markov chain and sojourn time in a state has its distribution depending on $X_n$ or $X_{n+1}$ or both.

Markov renewal processes were first studied by Pyke (1961(a), 1961(b)). Semi-Markov processes were first studied by Levy (1954) and Smith (1955). In stochastic models for many types of social behaviour these processes have been used (see Bartholomew (1967)). A semi-Markov model has been used by Kao (1974) in some context of hospital administration - in the study of dynamics of movement of patients through various care zones (e.g. medical unit, surgical unit, intensive care unit etc) in a hospital. The theory of birth process has been developed a Markov renewal process by Sheps (1964), (1973) and the theory has been applied to microdemography by Potter (1970).

We may note that if $Q_{1j}(t)$ in (1.5) has the form

$$Q_{1j}(t) = p_{1j} \left[ 1 - e^{-\lambda_i t} \right]$$

(1.8)

for all $i, j \in S$ and $t \in \mathbb{R}_+$ for some function $\lambda_i$, $i \in S$ then $Y = \{ Y(t), \ t \geq 0 \}$ will be a Markov process (cf. Cinlar (1975)).

In this thesis we consider a Markov renewal process in which change of state occurs in accordance with Markov chain with transition probabilities $p_{ij}$'s as defined in (1.6) and intertransition intervals
(t_i - t_{i-1}) (i = 1,2,..., n) have exponential distribution with parameter \( \lambda_i \) depending on state \( i \) only. In view of (1.7) here, we call such processes as continuous time Markov chain only as the time parameter is continuous and inter-transition intervals are exponential random variables. An additional property is that a transition (of \( \lambda_n \)) is allowed from a state into itself. It may be seen that this model differs slightly from the usual continuous time Markov chain model in which \( p_{1i} = 0 \). Here we assume that the state space \( S = \{1,2,...,m\} \) is finite containing no instantaneous states. We also assume that the process is regular and ergodic with no absorbing state possessing a stationary distribution. Thus it is possible to be in any state after some number \( N \) of steps and also to go from every state to every other state.

Let \( \lambda_{ij} = \lambda_i p_{ij} \ (i, j \in S) \). Then the hypothesis regarding \( \lambda_{ij} \)'s can be written as the hypothesis regarding \( \lambda_i \)'s and the hypothesis regarding \( p_{ij} \)'s. As we have mentioned earlier \( \lambda_i \) is the parameter of the waiting time. It may be seen that \( \lambda_{ij} \)'s may be defined as the rate of transition from state \( i \) to \( j \). The tests of hypotheses studied in this thesis actually pertain to this situation. However if we consider the Markov process \( Y(t) \) given in (1.7), then the rates \( \lambda_{ij} \)'s are defined for all \( i \neq j \), as in (1.2). We may apply the results in our thesis to this situation also, assuming suitable modifications of parameter.
It may be seen that if intertransition intervals \((t_1 - t_{i-1})\) 
\((i = 1, 2, \ldots, n)\) are (i.i.d.) random variables having exponential 
distribution with parameter \(\lambda\), then we have Markov chains subordinated 
to Poisson processes. (cf. Cinlar (1975), page 236). In addition, 
if \(X_0, X_1, \ldots, X_n\) are also i.i.d. random variables, then we have an 
independent increment process given by \(\{S_n, t_n\}\) where \(S_n\) is the 
sum of \(X_0, X_1, \ldots\) up to the \(n^{th}\) transition and \(t_n\) is the epoch of 
\(n^{th}\) jump.

In this thesis we study the asymptotic inference about Markov 
processes described earlier. We discuss the testing of hypothesis 
problem for testing the transition probabilities. In Chapters II and 
III we discuss the inference for Markov processes. In the last two 
chapters we consider non-parametric inference.

Asymptotic inference for discrete time Markov chains has 
been studied by many persons. The problems of testing of hypothesis 
about transition probabilities of a Markov chain has been considered 
by Anderson and Goodman (1957), Good (1955), Hoel (1954), Billingsley 
(1961), Goodman (1958) etc. Basawa and Prakasa Rao (1980) have also 
discussed this problem. Anderson and Goodman (1957) have obtained 
the likelihood ratio test and asymptotic chi-square tests for testing 
the hypothesis about transition probabilities of a Markov chain and 
also about the order of Markov chain. They have obtained the asymptotic 
results as \(n \to \infty\) where \(n\) is the total number of transitions.
Asymptotic inference about discrete Markov processes i.e. continuous time discrete state Markov processes has been discussed by many authors and to name a few of them here - Albert (1962), Allen (1983), Adke and Manjunath ((1984), (1984)(a)), Manjunath (1984), Billingsley (1961), Basawa and Prakasa Rao (1980) etc, among others. Albert (1962) has discussed estimation problem for transition intensities. He has obtained these results for finite Markov process also. He has obtained the asymptotic results as $T \to \infty$, where $T$ is the total time of observation. Adke and Manjunath (1984) have also considered the same hypothesis specifying the transition intensities for finite state Markov process and they have obtained the asymptotic distribution of the likelihood ratio test as $T \to \infty$.

In this thesis we obtain Neyman-Pearson likelihood ratio test and their asymptotic distribution under the null hypothesis, for testing the hypothesis about transition intensities $\lambda_{1j}$'s and also about transition probabilities $p_{1j}$'s of the Markov process (where $\lambda_{1j}$'s and $p_{1j}$'s are defined earlier). We use strong law of large number (SLLN) for i.i.d. random variables and also central limit theorem for random sums of i.i.d. random variables, to get the asymptotic properties of tests.

The likelihood function for discrete Markov process is given by

$$L = C \prod_{1 \leq i \leq m} \prod_{1 \neq j}^{n_{1j}} \lambda_{1j}^{n_{1j}} \prod_{1 \leq l \leq m} e^{-\lambda_{1j} Y_{1j}} (1.9)$$
where $\lambda_{ij}$'s and $\lambda_i$'s are as defined in (1.2) and (1.3),

\[ n_{ij} = \text{number of transitions from state } i \text{ to } j; \]
\[ n_i = \sum_{j=1}^{m} n_{ij} = \text{number of transitions from state } i \text{ to other possible states}; \]
\[ n = \sum_{i=1}^{m} n_i = \text{total number of transitions}; \]
\[ \gamma_1(k) = \text{total time spent in state } i \text{ when it visits for the } k\text{th time}; \]
\[ \gamma_1 = \sum_{k=1}^{m} \gamma_1(k) = \text{total time spent in state } i; \]
\[ T = \sum_{i=1}^{m} \gamma_1 = \text{total time of observational period}; \]

The likelihood function for the Markov process which we consider in this thesis is given by

\[ L = C \prod_{i=1}^{m} \lambda_{ij}^{n_{ij}} \prod_{i,j=1}^{m} e^{-\lambda_i \gamma_1} \quad (1.10) \]

where $\lambda_i = \sum_j \lambda_{ij} > 0$ are as defined in (1.3) for all $i$ and $j \in S$.

The likelihood is completely specified by $\lambda_{ij}$'s or $\lambda_i$'s and $p_{ij}$'s. In terms of $\lambda_i$'s and $p_{ij}$'s we have

\[ L = C \prod_{i=1}^{m} \lambda_i^{n_i} e^{-\lambda_i \gamma_1} \prod_{i,j=1}^{m} p_{ij}^{n_{ij}} \quad (1.11) \]

Thus we see that the likelihood function of Markov process can be split up into two factors, one depending on transition rates $\lambda_i$'s ($i = 1, 2, \ldots, m$) and the other on transition probabilities $p_{ij}$'s.
since \( p_{ij} = \frac{\lambda_{ij}}{\lambda_1} \) (\( i, j = 1, 2, \ldots, m \)). Thus hypotheses regarding \( \lambda_i \)'s can be tested independently of those regarding \( p_{ij} \)'s.

We also consider asymptotic measure of efficiency of the test statistic. Asymptotic measures of relative efficiencies can be broadly classified into two categories as (a) local and (b) non-local efficiency. A measure of performance that requires the alternative to approach the null is a local efficiency and a measure that lets the alternative stay fixed as \( n \to \infty \) where \( n \) is the sample size, is a non-local efficiency. Pitman power criterion is the popular local efficiency criterion. Basawa and Prakasa Rao (1980) have given another local efficiency criterion known as local power criterion. Most popular non-local efficiency is Bahadur efficiency in which for fixed alternative we consider the rate of convergence of the level attained by a statistic as \( n \to \infty \). Another non-local efficiency is defined by Hodges and Lehman (1956). In this thesis we discuss Bahadur efficiency of test statistics.

Bahadur efficiency of test statistics and estimates is discussed by Bahadur (1960), (1965), (1967), (1980), Bahadur, Zabel and Gupta (1980). For a sequence of i.i.d. random variables, the Neyman-Pearson likelihood ratio test statistic is shown to be optimal by Bahadur (1965) under some conditions. Raghavachari (1970) proved the same result without any conditions. Bahadur (1967) has proved the optimality of the likelihood ratio statistic for a sequence of independent random variables. Bahadur and Raghavachari (1971) have proved the optimality of the likelihood ratio statistic.
for a sequence of random variables and they have discussed the examples of Markov chain also. Bahadur (1967) has defined the approximate slope of a test statistic. He has shown that, generally, exact slope and approximate slope will not be the same. But for the likelihood ratio statistic these two slopes are the same. We have given the definition of the exact slope and approximate slope in Chapter IV. Berk and Brown (1978) have discussed the sequential Bahadur efficiency in which case they study the rate of convergence of the level attained test statistic as the stopping time n goes to infinity. In Chapter IV of this thesis we show that the likelihood ratio tests obtained in Chapters II and III are Bahadur optimal.

Lastly in Chapter V we propose a non-parametric test for testing the hypothesis that the observed realization is a Markov process against the alternative that it is a semi-Markov process. We obtain the results on lines similar to those of Proschan and Pyke (1967). In Chapter VI we propose a non parametric test to test the hypothesis that two realizations of Markov process are from the same Markov process. The results are obtained on the lines similar to that of Mann and Witney (1947). In both the chapters these nonparametric tests are shown to be consistent and unbiased.

2. Schemes of observation of the process

Markov process can be observed as per the sampling schemes. In this thesis we obtain the asymptotic results for the process observed under both schemes.
Scheme 1. The Markov process is observed continuously up to the $n$th transition.

Scheme 2. The Markov process is observed continuously over a period $(0, T)$.

Thus under scheme 1, the total number of transitions $n$ is fixed and the time of observation $T$ is a r.v. Under scheme 2, $n$ is a r.v. and the total time of observation $T$ is fixed. When the process is observed under scheme 1 we obtain the asymptotic results as $n \to \infty$ and when it is observed under scheme 2, we obtain the asymptotic results as $T \to \infty$. Both the sampling schemes assume the availability of complete information on the trajectory of the process. Billingsley (1961) has obtained the likelihood ratio tests and (asymptotic) chi-square tests for testing the hypotheses about transition intensities of Markov process observed under scheme 2. Adke and Manjunath (1984) have also discussed the inference problems for finite Markov process observed under scheme 2. Basawa and Prakasa Rao (1980) have studied homogeneous Poisson process observed under both the schemes 1 and 2. They have discussed the estimation and also testing of hypothesis problem. Albert (1962) has also studied the Markov chain with continuous time parameter observed under scheme 2. Basawa (1974) has considered the problem of estimation for renewal and Markov renewal process under these two sampling schemes. Moore and Pyke (1968) have discussed the estimation of transition distributions of Markov renewal process observed under scheme 2. Moran (1951) has discussed the estimation of parameters for pure birth process observed under scheme 1 which
he defined as inverse sampling plan. Moran (1953) has discussed the estimation of parameters of birth and death processes observed under scheme 1.

In this thesis we consider a Markov process as described earlier, under both schemes 1 and 2 and obtain the asymptotic results.

3. Chapterwise Summary

This thesis has six chapters. Chapter II considers certain hypotheses about continuous time finite state Markov chain. As mentioned earlier, a Markov process is completely specified by \( \lambda_{ij} \)'s or \( \lambda_1 \)'s and \( p_{ij} \)'s (\( i, j = 1, 2, \ldots, m \)), where \( \lambda_{ij} \)'s and \( p_{ij} \)'s are as defined earlier at \( \text{(2.1.3)} \). The likelihood function of Markov process can be split up into two factors one depending on \( \lambda \)'s and the other depending on \( p_{ij} \)'s. The hypothesis regarding \( \lambda \)'s can be tested independently of those regarding \( p_{ij} \)'s. We obtain the likelihood ratio tests for testing these hypotheses and show that they have asymptotic chi-square distribution under the null hypothesis as \( n \to \infty \) when the process is observed under scheme 1 and as \( T \to \infty \) when the process is observed under scheme 2. We use central limit theorem for random sums of i.i.d random variables to get the asymptotic distribution of these tests. It is noted that the likelihood ratio test for testing the hypothesis about both \( \lambda \)'s and \( p_{ij} \)'s is the product of the likelihood ratio test for testing \( \lambda \)'s and \( p_{ij} \)'s respectively. Since \( \lambda \)'s and \( p_{ij} \)'s can be tested independently we note that the likelihood ratio test for testing the hypothesis about \( p_{ij} \)'s is the
same as that of discrete time Markov chain as discussed by Anderson and Goodman (1957) and others. The hypothesis specified in Theorem (2.2.4) is the same as testing the hypothesis that $X_1, X_2, \ldots, X_n$ are Markov chains subordinated to Poisson processes (Cf. Cinlar (1975), page 236) against the alternative that $(X, T) = \{X_n, t_n ; n \in \mathbb{R}_+\}$ is a Markov process. The hypothesis specified in Theorem (2.2.6) is the same as testing the hypothesis that $S_n = \sum_{i=1}^{n} X_i$ is an independent increment process against the alternative that $(X, T)$ is a Markov process. This hypothesis will enable us to test the hypothesis that the Markov process is of order zero against the alternative that it is of order one. We also obtain the likelihood ratio test for testing the homogeneity of several samples of Markov processes. We consider three hypotheses, each of which can be interpreted as a test for homogeneity of several samples. These likelihood ratio tests have asymptotic chi-square distribution under the null hypothesis. These tests have no parallel in case of discrete time Markov chains.

In Chapter III we define a Markov process of order $r$. We obtain likelihood ratio tests for testing the hypothesis that the Markov chain is of order zero against the alternative that it is of order one. We consider three hypotheses, each of which can be interpreted as a test of hypothesis that the Markov chain is of order zero against the alternative that it is of order one. These tests also have no parallel in case of discrete time Markov chains. Similarly we obtain the likelihood ratio tests for testing the hypothesis that the Markov chain is of order one against the alternative that it is
of order two. We consider these two special cases for the clarification of thoughts and their practical importance per se. The likelihood ratio test for testing the hypothesis that the Markov chain is of order r-1, against the alternative that it is of order r is obtained and also for testing the hypothesis that the Markov chain is of order u (u < r) against the alternative that it is of order r. All of these tests are shown to have asymptotic chi-square distribution under the null hypothesis, as \( n \to \infty \) when the process is observed under scheme 1 and as \( T \to \infty \) when the process is observed under scheme 2.

In Chapter IV we obtain the exact slopes defined by Bahadur (1960) and approximate slopes defined by Bahadur (1967) of all the likelihood ratio tests obtained in Chapter II and Chapter III and show that they are Bahadur optimal in the sense of slopes. We see that the generalized Kullback-Leibler number for testing the hypotheses about \( \lambda_1 \)'s and \( p_{ij} \)'s is the sum of the Kullback-Leibler numbers for testing the hypotheses about \( \lambda_1 \)'s and \( p_{ij} \)'s independently, and also the slopes of likelihood ratio tests for testing \( \lambda_1 \)'s and \( p_{ij} \)'s is the sum of the slopes of likelihood ratio tests for testing \( \lambda_1 \)'s and \( p_{ij} \)'s respectively. The exact and approximate slopes of these likelihood ratio tests are shown to be the same. It may be noted that the slope (exact as well as approximate) of these likelihood ratio tests for testing the hypotheses about Markov process observed under scheme 1 is the same as the slope (exact and approximate) of tests for Markov process observed under scheme 2. Thus these likelihood ratio tests for Markov process are equally efficient in the sense of slope when
it is observed under both schemes 1 and 2.

In Chapter V we propose a nonparametric test for testing the hypothesis that the observed realization is a Markov process against the alternative that it is a semi-Markov process with the transition rate monotone increasing (or decreasing). The test criterion is shown to have asymptotic normal distribution under the null hypothesis. The test is shown to be consistent and unbiased.

In Chapter VI we propose a nonparametric test for testing the hypothesis that two realizations of a process are from the same semi-Markov process. This test criterion is shown to have asymptotic normal distribution under the null hypothesis and the test is shown to be consistent and unbiased.

4. Scheme of reference to equations and sections.

The equations and sections within a chapter are referred to simply by their numbers in the same chapter. Thus in Chapter II the equation (2.5) is referred to as such only. But outside the chapter equation (2.2.5) is referred to as the equation (2.5) of Chapter II. Similarly Theorem (2.6) of Chapter II is referred to as Theorem (2.2.6). We use \( o(1/n) \) as a symbol to denote a function which tends to zero as \( n \to \infty \).