CHAPTER I
INTRODUCTION

1.0 Introduction

The theory of stochastic process has developed considerably in the last three decades. Its field of application is constantly expanding and at present it is being applied nearly every branch of science and technology. The theory of stochastic process plays an important role in the investigation of random phenomena depending on time. The earlier results in this direction were concerned with the investigation of Brownian motion, telephone traffic, shot noise of electronic tubes, etc. A. Einstein (1912), M. Smoluchowski (1906), L. Bachelier (1912), A.K. Erlang (1918), W. Schotlky (1918). The foundation of mathematical theory of stochastic process was given by A.N. Kolmogorov (1931). Since that time the theory and applications of stochastic processes have shown constant development.

The statistical inference in stochastic process is gaining much momentum in these days. Today several basic problems have been solved, resulting in a vast literature, in part motivated by applications, among which communication engineering should be mentioned first. The theory is still actively developing, as can be seen from the articles in the leading journals of mathematical statistics and probability.
The classical theory of estimation and of testing hypothesis has been already extended by many authors to study some problems of stochastic processes. Mainly the work is based on construction of the likelihood which to be used either in estimating the parameters or in constructing a test of hypothesis. As usual for simple likelihood, the methods are quite elegant and wherever the likelihoods are of difficult form, then the asymptotic theory has been developed. Standard properties wherever possible of maximum likelihood estimators have been verified for the class of stochastic processes. Some new properties are also found.

Much of the standard inference methodology can be said to be likelihood based. A first step in solving the problem of inference is the construction of the likelihood function. This is obviously so for maximum likelihood estimation, for Neyman Pearson approach of testing of hypothesis and for the notion of confidence regions. Actually the same holds for Bayesian inference also.

Not much work has been done to see the possibility of extending the classical theory of sequential inference problems to stochastic processes. This thesis is mainly devoted to develop some methods of sequential inference for the parameters of a certain class of stochastic processes.
The sequential procedure is a method of statistical inference whose characteristic feature is that the number of observations required or the time required for observing a process is not determined in advance. The decision to terminate the observations on the process at each stage depends on the results of the observations previously made. A merit of the sequential method is that, test procedures and estimation procedures can be derived which require on the average, a substantially smaller duration of observation or smaller number of observations than equally reliable procedures based on a predetermined duration of observation or number of observations, respectively.

A general problem of sequential estimation suggests itself, namely, to formulate a rule of sampling such that an unknown population parameter can be estimated with specified accuracy and with minimum expected sample size, whatever be the true value of the parameter. The precision of estimation might be specified by width of the confidence interval or by coefficient of variation of the estimate or in some other way. One such estimation problem is proposed by Haldane (1945). He considered a sequential method of sampling in which sampling terminates when a specified number \( k \) of individuals have been found possessing a certain attribute. Wald (1947) in his 'Sequential Analysis' devoted a short section to discuss sequential
procedures whose primary objective was to provide a confidence interval for an unknown parameter having preassigned width or satisfying some other similar condition with as few observations as possible.

The problem of minimax sequential estimation with a special form of cost function has been considered by Wald (1951). Cox (1952) has given a method for finding, at the end of the sequential sampling procedure, a nearly unbiased estimate of the population mean and derived the bias of Anscombe (1949) estimate when the unknown parameter is the population mean. All the above problem have been solved by two-stage sampling technique.

1.1 Two-stage Sampling

The two-stage sampling method was first introduced by Dodge and Romig (1929) and was later investigated by Stein (1945). In this two-stage sampling scheme the decision whether a second sample should be drawn or not depends on the outcome of the observations in the first sample. The reason for introducing a two-stage sampling method was, of course, the recognition of the fact that the two-stage sampling results in a reduction of amount of inspection as compared with single stage sampling. The two-stage sampling method does not fully take advantage of sequential analysis, since it does not allow for more than two-samples. A
multiple sampling scheme results in a reduction of amount of inspection as compared with two-stage sampling.

Any two-stage sample procedure, such as that discussed in the paper by Steins (1945) can be considered a special case of sequential analysis developed by Wald (1945). Stein (1945) presented a two-stage sample test for linear hypothesis, in which the size of the second sample depends upon the results of the first. The problem determining confidence intervals of pre-assigned length and confidence coefficients for the mean of a normal distribution with known variance has been discussed. Stein (1945) remarks that the problem of whether the two stage sample tests and confidence intervals are in any sense optimum is unsolved. However he has shown that, if the variance and initial sample sizes are sufficiently large, the expected number of observations differs only slightly from the number of observations required for a single-sample test when the variance is known. Many other authors have discussed the sequential estimation problems based on two-stage sampling technique. Namely Cox (1952), Birnbaum and Healy (1960), Samuel (1966), etc.

1.2. Sequential Estimation

Estimation by two-stage sampling method has been discussed by Cox (1952). He considered the problem of
estimating the unknown parameter $\theta$ with pre-assigned accuracy using as few observations as possible. For any parameter $\theta$ the mean sample size for the two-stage sampling solution will be greater than the corresponding mean sample size for the best sequential procedure. However, Cox (1952) has shown that the difference is likely to be small except when the sample sizes are very small. The basic idea is familiar (see for eq: Quenouille, 1950) — a preliminary sample is used to determine how large the total sample should be.

In most of the problems of estimation, estimators based on samples of fixed sizes have precisions which depend on unknown parameters; and estimators with prescribed precisions are not available without resort to sequential sampling in two or more stages, as in Stein's (1945) procedure for estimation of the mean of a normal distribution with unknown variance. However, the two-stage sampling method discussed by Cox (1952, Vol. 40, 39) and the sequential method of Anscombe (1953) seem to be both fairly practicable and efficient. The latter methods, however, approximate being based on asymptotic theory, and there seems to be no method which is easily applicable for estimation.
An approach employing a different concept of 'precision' is described by Birnbaum and Healy (1960). They considered the problem of estimating a real valued function \( f(\theta) \) of an unknown parameter \( \theta \) with prescribed precision; and they suggested an unbiased estimator \( \hat{\theta} \) having variance not exceeding a given positive function \( B(\theta) \), which depends on the parameter \( \theta \). The method described by Birnbaum and Healy (1960) is simple one. It provides in a number of problems, procedures for two-state sampling leading to estimators which are unbiased; in certain problems these estimators have a prescribed variance, while in other problems they have variances never exceeding but generally close to a prescribed bound. The estimation methods of Birnbaum and Healy (1960) are almost similar to the method of Stein (1945) for estimation of a normal mean.

On similar lines with little modifications Samuel (1966) considered the problems of unbiased estimation of the parameter \( \theta \) of binomial and Poisson distributions. The variance of these estimators are required not to exceed a pre-assigned bound \( B \) which is independent of the parameter \( \theta \). The method discussed by Samuel (1966) differs from the method Birnbaum and Healy (1960) in utilising the information contained in the preliminary sample. Samuel (1966) used the first sample not only to determine the size of the
second sample, but also in estimating the parameter $\theta$; whereas Birnbaum and Healy (1960) used the second sample information to estimate the parameter $\theta$. Even though the method discussed by Samuel (1966) is shown to be better in the sense of saving in number of observations in some particular cases, he has not claimed any optimal properties.

The next two chapters of this thesis deals with the problem of sequential estimation of a parameter $\theta$ of certain continuous time stochastic processes based on two-phase observation technique. We have obtained an unbiased estimate of the parameter $\theta$. The information available through the first phase observation will be used to determine the length of the period of the observations at the second phase; and for estimating the parameter $\theta$ we use both the phases of observations. Some interesting properties and inequalities have been derived.

1.3 Closed sequential decision procedures

Several authors (viz; Armitage and Schneiderman (1962), Dworetzky, Kiefer and Wolfowitz (1953), etc) have discussed sequential decision procedures. Wald (1947) started a programme for dealing with multiple decision problems based on sequential estimation. This programme was never completed and it was to some extent further
developed by Stein (1948).

Sobel and Wald (1949) considered a problem of a sequential decision procedures for choosing one of three hypothesis concerning the unknown mean of a normal distribution. Armitage and Schneideman (1962) considered a particular case of Sobel and Wald (1949) general formulation. Further they have remarked that, Sobel and Wald (1949) have given upper and lower bounds for the ASN but these bounds are occasionally wide, and since over shoot of the boundaries is neglected, their values may not be true bounds. Hall (1962) presents several sequential analogues of Steins two-stage test procedure for testing hypothesis about the mean of a normal population with unknown variance and with specified bounds on the error probabilities.

Using different approach, Paulson (1963) has given a sequential procedure for deciding as to which of $k$ non-overlapping intervals the unknown mean $\theta$ belongs. The procedure is to satisfy the requirement that the probability of making an incorrect decision is less than some pre-assigned value $\alpha$. The sequential procedure is worked out explicitely for two cases, namely, one is when $\theta$ is the mean of a normal distribution with a known variance and another is when $\theta$ is the mean of a normal distribution with an unknown variance. In his brief discussion a
related but apparently new problem to find a sequential
procedure is given which will simultaneously select one of
k intervals and also yield a confidence interval for $\theta$ of
specified width. However Paulson (1963) does not claim any
optimal properties of his solution, but nevertheless it is
hoped that the results may be useful in practice.

Using slight modification of a procedure given by
Wald (1947, chapter 10) Paulson (1964) obtained the sequential
confidence limits for the parameter of a normal distribution.
He obtains a new derivation of sequential confidence
intervals for the mean, for the variance and for the ratio
of the variances of two normal distributions. Also the
results are applied to get closed decision procedures for a
number of decision problems including the test of hypothesis
regarding the mean, variance and ratio of variances of
normal distributions against either one sided or two-sided
alternatives. In addition a brief discussion of 'mixed'
problems concerned with finding a confidence interval for
the parameter after a decision has been reached is discussed.

The closed test has the advantage that for small
values of $\alpha$ and $\beta$ it substantially reduces the maximum
average sample size. In these respect Paulson (1964) test
is somewhat similar to a modification of the Wald's
sequential test given by Anderson (1960), when $\sigma$ is known.
One more advantage of Paulson's (1964) procedure is that it can be easily modified to solve the problem when $c'$ is unknown. Closed sequential procedures were given by Armitage and Schneider (1962) for nine different combinations of $\alpha$ and $\beta$. Hall (1962) also gave a closed sequential solution for the special case $\alpha = \beta$, which is similar to but not identical with the solution given by Paulson (1964). Paulson's (1964) procedure seems to compare reasonably well with the sequential probability ratio tests of Wald.

The IV-th chapter of this thesis deals with finding the confidence interval and closed sequential procedures for the parameter of Koopman-Darmois family of processes. The last chapter deals with the sequential procedures for life-testing.

1.4 Chapterwise Summary

In the II-nd chapter for the Koopman-Darmois family of processes indexed by a real parameter $\theta \in (a, b)$, the problem of unbiased estimation of $\theta$ when the variance of the estimate is required not to exceed a pre-assigned upper bound is discussed. Using the transitive sufficient statistics, a method of achieving this through a sequential method has been suggested. Here we consider two-phase observation procedure. The first-phase observation is not
only used to determine the period of second-phase observation, but is also directly used in estimating the parameter $\theta$.

The III-rd chapter is concerned with minimum variance unbiased estimation subject to an upper bound on the variance of estimators, using as minimum period of observation as possible. In this chapter we have obtained an unbiased estimate of the parameter, by different method, for the Koopman-Darmois family of processes via two-phase technique. The method has been illustrated by some particular examples.

In the IV-the chapter, the closed sequential decision procedure suggested by Paulson (1964) has been extended to a general family of distribution. The method of obtaining the approximate formula for the Average Sample Number (ASN) has been given. In particular the formula for ASN when testing a hypothesis regarding the mean and variance of the normal distribution has been derived. Also we have obtained the sequential confidence limits for the mean of the Koopman-Darmois family of processes. And we have obtained a closed sequential procedure for AST function of Koopman-Darmois family of processes. The method has been illustrated by some particular examples.
The V-th chapter deals with the closed sequential procedure for life-testing problems. Suppose there are n independent components or units with identically distributed failure-times \( T_1, T_2, \ldots, T_j \) having an exponential failure-time distribution, which depends on an unknown parameter \( \theta \); we have obtained the sequential confidence interval for the reliability function \( R(t/\theta) \). A closed sequential decision procedure for testing of hypothesis regarding \( R(t/\theta) \) has been suggested. An approximate formula for \( \Delta \text{SN} \) for this procedure is given.