CHAPTER 1

INTRODUCTION
INTRODUCTION

1.1 INTRODUCTION AND REVIEW OF LITERATURE

The problems of estimation of parameters and prediction of future random variables are of fundamental significance to any statistical investigation. The theory and techniques of estimation and prediction occupy the prime position in supporting scientific investigation in many fields.

This thesis is a compilation of the work done by the author in the fields of estimating functions and prediction.

Conventionally the problem of estimation is tackled by obtaining an estimate through some standard methods such as least squares, maximum likelihood, minimum chi-square and method of moments. A common feature of these methods is that they lead to 'estimating equations'. The phrase 'equation for estimation' was used more than half-a-century ago by Fisher (1935) and the now famous term 'estimating function' was given by Kimball (1946). The theory of estimating functions combines the strengths of maximum likelihood and least squares methods and that of minimum variance unbiased estimation and at the same time eliminates some of their weaknesses and presents a unified approach to estimation problems (Refer Godambe and Kale (1991)).

While the likelihood method of estimation was intuitively understood and already in vogue, it was not based on any 'optimality' consideration until Godambe (1960) established that under certain regularity conditions on the class of densities and the class of unbiased estimating functions (UEFs) the score function is optimal for models involving a single unknown parameter. Kale (1962) derived an extension of Cramer-Rao inequality for UEFs. It was primarily owing to the pioneering
contributions of Godambe (1960) and Kale (1962) that the theory of estimating functions (EFs) gained momentum and attracted the attention of many authors. Later, Bhapkar (1972) produced a similar result for vector parameter situation in a different approach. These authors gave a formal justification for the likelihood method in small samples for parametric families when all the indexing parameters are of interest.

Godambe and Thompson (1974) introduced unbiased estimating functions and an optimality criterion for the problem of estimating a single interesting parameter in the presence of a nuisance parameter. Godambe (1976) introduced standardised form of UEFs for the same problem in the presence of many nuisance parameters and proved the essential uniqueness of optimal UEFs.

In order to get optimal UEF, various optimality criteria have been proposed and applied in the literature. Ferreira (1982 a) considered the case of vector interesting parameter and proposed an optimality criterion for vector UEFs based on non-negative definiteness of difference of dispersion matrices. Chandrasekar and Kale (1984) considered three optimality criteria : (i) M-optimality, same as that considered by Ferreira (1982 a), (ii) T-optimality, based on trace of dispersion matrices, (iii) D-optimality, based on determinant of dispersion matrices, and established their equivalence assuming that an M-optimal UEF exists. Later, Chandrasekar (1988) proposed $Q_A$-optimality based on quadratic forms and established that this is also equivalent to the three former criteria, again assuming the existence of M-optimal UEF. Joseph and Durairajan (1991) established the equivalence of these four criteria without assuming the existence of M-optimal UEF following the characterisation of optimal UEF obtained by Durairajan (1989). Recently, Yanagimoto and Yamamoto (1993) extended the criterion of Godambe and Thompson (1974) for the case of real interesting parameter with nuisance parameters
to measure the efficiency of biased EFs also. Their criterion is, however, not applicable to semi-parametric models. Hitherto, optimality criterion for biased estimating functions applicable to both parametric and semi-parametric families has not been proposed. An attempt is made in this direction in the thesis.

In the theory of UEFs applied to parametric models involving nuisance parameters, the 'elimination' of nuisance parameters to obtain optimal EF for interesting parameters is a very important task. Godambe (1976) suggested a method of eliminating nuisance parameters by multiplying and adding suitable functions to the score function. Formally, he assumed a 'conditional' factorisation property and under different sets of assumptions established the optimality of conditional score function, which is the derivative of the logarithm of the conditional density of the data given an appropriate statistic. Lloyd (1987) and Bhapkar and Srinivasan (1993) assumed a 'marginal' factorisation property and claimed the optimality of 'marginal score function' which is the derivative of the logarithm of the marginal density of an appropriate statistic. However, that there are certain gaps in the results of Lloyd (1987) and Bhapkar and Srinivasan (1993) has been discovered in a recent paper by Bhapkar (1995) who imposed some more conditions and established the optimality of marginal score function. These 'factorisation' properties are of much use in nuisance parameter elimination. A unified method or approach for elimination of nuisance parameters without going into these factorisation aspects does not seem to have been developed so far and this gap is sought to be filled up in this thesis.

Just like other branches of statistical inference, the theory of EFs also has been approached from the Bayesian viewpoint. Ferreira (1982 b) initiated a Bayesian approach by suggesting integration with respect to the data and the parameter. He established the optimality of the modal posterior function within a certain class of
EFs. Ghosh (1990) suggested another approach by carrying out integration only over the parameter space conditional on the data in hand. This approach, as Ghosh has pointed out, is more in conformity with the Bayesian paradigm. By imposing some regularity conditions on the model and on the class of posterior unbiased estimating functions he also established the optimality of modal posterior function. In this thesis following the approach of Ghosh, we additionally consider posterior biased EFs under weaker assumptions on the model.

An important breakthrough in estimating function theory in the context of estimation for stochastic processes under semi-parametric model was achieved by Godambe (1985, 1987). While generally in the literature on stochastic processes estimation is investigated in terms of asymptotic properties, the theory put forward by Godambe (1985, 1987) led to finite sample optimal estimation. Subsequently, Godambe and Thompson (1989) extended the concept and technique of Quasi-likelihood estimation suggested by Wedderburn (1974) and developed the extended quasi-likelihood function for independent observations by incorporating knowledge of skewness and kurtosis. The quasi-score function has thus been identified with optimal estimating function and a new and fruitful perspective on quasi-likelihood has been provided by the theory of estimating functions. In all the above-mentioned works of Godambe (1985, 1987) and Godambe and Thompson (1989) the UEFs are constructed as linear combinations of a basis of orthogonal EFs. Durairajan (1992) considered linear combinations of a basis of non-orthogonal EFs and obtained explicit closed form for optimal UEF. Using this approach, Bai and Durairajan (1996) obtained finite sample optimal estimation for the parameters of branching process with immigration and two-type branching process. Using the same approach of Durairajan (1992), we develop a generalised quasi-likelihood estimation for dependent processes with martingale structure incorporating knowledge of higher order moments.
Another important branch of statistics which is most relevant in many practical situations is prediction theory. As Ishii (1969) stated, one of the ultimate aims of Statistics is to declare something about future random variables. Some of the early works on prediction theory are those of Goldberger (1962) on best linear unbiased prediction in the Generalised linear regression model, Hora and Buehler (1967) on invariant prediction and Ishii (1969) on unbiased prediction.

The theory of equivariant estimation and prediction are appealing in the sense of obtaining estimators and predictors with nice mathematical and operational properties when symmetries are present in a problem. For equivariant estimation in location, scale and location-scale families we refer to Ferguson (1967) and Lehmann (1983) and for equivariant prediction in location-scale families we refer to Kaminsky, Mann and Nelson (1975), Takada (1981) and Natarajan and Durairajjan (1991). In this thesis simultaneous equivariant vector prediction is considered for location, scale and location-scale families and its equivalence to component-wise equivariant prediction is established.

The theory of equivariant prediction has also been applied fruitfully in the context of finite population sampling and a few interesting contributions exist in the literature. Prediction of population variance is considered in Zacks (1981), Zacks and Solomon (1981) and Zacks and Bolfarine (1991). Bolfarine (1989) considered equivariant prediction of population total under various super-population models with respect to squared error loss. In the finite population context, the prediction of the population total with respect to a general even convex loss function does not seem to have been attempted in the literature and this is achieved in the thesis.

As reviewed above there is a considerable literature on unbiased and equivariant prediction in the context of infinite and finite populations. Prediction in
stochastic processes has been considered by Yamamoto (1976), Fuller and Hasza (1981), Basu and Sen Roy (1989) among others with stress on asymptotic properties. However, prediction through estimating functions was not thought of until Durairajan (1996) invoked UEFs for simultaneous estimation and prediction in discrete-time stochastic process under a semi-parametric model with stress on finite sample optimal property. Durairajan (1996) considered the problem of predicting one future random variable in addition to estimating the interesting parameters. The prediction of a vector of future random variables in this approach is taken up in the thesis.


A major part of the thesis focuses on various facets of estimating functions and constitutes an attempt to contribute some notions, results and techniques to the theory of estimating functions. The thesis also attempts optimal prediction in semi-parametric models, group families and finite population sampling.
1.2 OUTLINE OF THE THESIS

The thesis is organised as follows. In each of the chapters of the thesis, the First Section gives the motivation for and a brief introduction to the chapter.

Chapter 2 focuses on a wide class of EFs which includes both biased and unbiased EFs. In Section 2.2 a new criterion is defined and justified. Section 2.3 provides the extensions of Rao-Blackwell theorem and Cramer-Rao inequality with reference to this criterion. A characterisation of optimal EF along the lines of Durairajan (1989) is obtained in Section 2.4. In Section 2.5, invoking the 'invariance' principle, the notion of invariant estimating function is introduced in the context of parametric families which are invariant in a specific sense and examples are discussed. The applicability of the criterion to semi-parametric models is demonstrated in Section 2.6.

In Chapter 3 the notion of weakly unbiased estimating function is introduced in the context of parametric families with nuisance parameters and with reference to the new optimality criterion proposed in Chapter 2. In Section 3.2 a sequence of Bhattacharyya type lower bounds for the mean square error matrix of weakly unbiased EFs is derived and explicit forms of the functions attaining the bounds are obtained. The results presented here suggest a direct and unified procedure for possible elimination of nuisance parameters without going into the factorisation (conditional / marginal ) aspects of the density. Examples are discussed in Section 3.3 wherein some of the existing examples in which the claim on optimal UEF are incorrect are resolved in a methodical approach. The highlight is the provision of a common recipe for deriving optimal EFs in an 'iterative' manner. Some comments on the results presented here and related results are given in Section 3.4.
Chapter 4 discusses estimating functions in the Bayesian paradigm as suggested by Ghosh (1990). In Section 4.2 the optimality of the modal posterior function is established under weaker assumptions on the model in a wider class of EFs that contains the class considered by Ghosh (1990). In Section 4.3 the nuisance parameter situation is considered. Section 4.4 discusses the multiparameter situation wherein the main result is proved through characterisation of optimal EF as distinct from the inequality approach adopted in Section 4.2.

Chapter 5 develops a generalised quasi-likelihood estimation for discrete-time processes whose dependent structure satisfies some prescriptions and for which knowledge on skewness and kurtosis are available. In Section 5.2 we review Durairajan (1992) whose approach is invoked for developing the generalised quasi-likelihood estimation. In Section 5.3 the semi-parametric model is described, the required prescriptions are laid down and the generalised quasi-score function is derived. Applications are discussed in Section 5.4 and some comments on the results are given in Section 5.5.

In Chapter 6 optimal estimating function in some specific semi-parametric models are derived. In Section 6.2 an exchangeable location model is considered and the optimal UEF is derived. It is shown that the UEF derived in Godambe and Thompson (1978) for the i.i.d. case is robust in the sense that it is optimal for the wider exchangeable location family which contains the family considered by Godambe and Thompson (1978). In Section 6.3 a bivariate model and a symmetric model are considered and the optimal UEF among linear combinations of certain EFs are derived invoking the theory of Godambe (1985, 1987) and Durairajan (1992).

Chapter 7 applies the theory of EFs for prediction of a vector of future random variables in semi-parametric models and extends the approach of
Durairajan (1996). Section 7.2 gives the extended versions of the results in Durairajan (1996). In the subsequent sections prediction for linear model, auto regressive process, branching process with and without immigration and two-type branching process are discussed in the above framework.

Chapter 8 deals with vector prediction in group families with respect to matrix version of squared error loss. The equivalence of joint optimality of vector predictor and marginal optimality of component predictors is established in each of location, scale and location-scale families.

Chapter 9 addresses equivariant prediction of a finite population total under a regression super-population model. Section 9.2 gives some basic definitions and results. In Section 9.3 a characterisation of minimum risk equivariant predictor (MREP) with respect to an even convex loss function is obtained and illustrations are given. While Bolfarine (1989) assumed an independent structure and derived MREP with respect to squared error loss in the presence of one auxiliary variable by exploring the relationship with best unbiased predictor (BUP), here, in contrast, the MREP is found by allowing a dependent structure and with respect to even convex loss function in the presence of many auxiliary variables and in a direct manner without reference to BUP.

To summarise, the salient features of this thesis are as follows:

1. An optimality criterion for biased estimating functions applicable to parametric as well as semi-parametric models is introduced.

2. A notion of invariant estimating function applicable to families of distributions that are invariant is proposed.