CHAPTER- 1

INTRODUCTION

1.1 INTRODUCTION :

Statistical inference is the process by which information from sample data is used to draw conclusions about the population from which the sample was selected. The techniques of statistical inference can be divided into two major areas, i) the problem of estimation and ii) the problem of testing of hypothesis.

The estimation problems which we consider are those called as parametric. The data consists of observation $X_1, X_2, \ldots, X_n$ whose distribution F_θ, $\theta \in \Theta$ has the known form except that it contains an uknown parameter θ. Further assume that the sample values $x_1, x_2, \ldots, x_n$ of $X_1, X_2, \ldots, X_n$ can be observed. On the basis of the observed sample values, it is desired to estimate the value of the unknown parameter θ or value of some function, say $\tau(\theta)$ of the unknown parameter. This process is called as 'Estimation'. That is, to let the value of some statistic $t = t(X_1, X_2, \ldots, X_n)$, estimate the unknown function of parameter $\tau(\theta)$; such a statistic is called an 'estimator' and the method is called estimation.

Further, the problem of estimation can be classified into two categories, viz, i) problem of point estimation and ii) interval estimation. A point estimate of a population parameter is a single numerical value of a statistic which corresponds to the unknown parameter. That is, the point estimate is a unique selection for the value of an unknown parameter. More precisely, if X is a random variable with probability distribution f(x), characterized by the unknown parameter θ, and $X_1, X_2, \ldots, X_n$ is a random sample of size n from X, then the statistic, $\hat{\theta} = h(X_1, X_2, \ldots, X_n)$ corresponding to θ is called the estimator of θ. Note that the estimate $\hat{\theta}$ is a random variable, beceause it is a function of sample data. After the sample has been selected $\hat{\theta}$ takes on a particular numerical value called the point estimate of θ.

In many situations, a point estimate does not provide enough information about the parameter of interest, but we can specify a random interval for unknown parameter, that is first find two statistics L(X) and U(X) such that

$$P\left[L(X) \le \theta \le U(X)\right] = 1-\alpha.$$

The resulting interval $L(X) \le \theta \le U(X)$ is called a $100(1-\alpha)\%$ 'confidence interval' for the unknown parameter θ. L(X) and U(X) are called the lower and upper confidence limits, respectively, and $1-\alpha$ is called the confidence coefficient.

The second problem is of testing of hypothesis. Let $X_1, X_2, \ldots, X_n$ be a sample from a distribution F_θ ; $\theta \in \Theta$ where the functional form of F_θ is known except that it contains an unknown parameter θ. A hypothesis is a statement or claim about the state of nature which needs verification, that is either rejection or acceptance for one or the other purpose. Then we are interested in testing the validity of the hypothesis. This process is called 'testing of hypothesis'. While formulating testing of hypothesis problem, we formulate the null hypothesis H_o and alternative hypothes is.

Now our interest is to focus on estimation only. The theory of estimation had taken a gaint leap forward. The problem of finding minimum variance unbiased estimator for the parametric function of the interest has got importance in statistical inference. The concept of unbiasedness as 'lack of systematic error ' in the estimators was introducsed by Gauss(1821) in his work on the theory of least squares. There can be many unbiased estimators for the same parametric functions. To choose one among these, one should have some criteria. One of the possible criteria is based on variance. Usually based on this criteria we choose an estimator, having the least variance. This leads to the concept of minimum variance unbiased estimator.

The first MVU estimator was obtained by Aitken and Silverstone(1942), in the situation in which the information inequality yields the same result. MVU estimators as unique unbiased functions of a suitable sufficient statistic were derived in special case by Halmos(1942) and Kolmogrov(1950). This was pointed out as a general fact by Rao(1947). Early use of unique unbiased functions of a suitable complete sufficient statistic is due to Tweedie(1947).The concept of completeness was defined, its implications for unbiased estimation developed and theorem is obtained, by Lehmann and Sheffe (1950, 1955, 1956).

There are virtually thousands of papers in the statistical literature on various topics of statistical inference. Results of many of these interesting papers can be found in Zack(1970) and Lehmann(1983). In the present thesis, we have concentrated on minimum variance unbiased estimation of parametric functions in non-regular families, namely one-truncation and two-truncation parameter families. Methods of deriving minimum variance unbiased estimators for one-truncation parameter families, were developed in several different studies and unified in general framework by Tate(1959). In this case the density functions are generally of the form,

$$f_1(x;\theta) = q_1(\theta)\ h_1(x)\ I_{(\theta,b)}(x)$$

or

$$f_2(x;\theta) = q_2(\theta)\ h_2(x)\ I_{(a,\theta)}(x).$$

Where $I_A(x)$ is the indicator function of set A. In other words, the location parameter is a point θ (unknown), which is at the left or right limit of the support interval of $f(x;\theta)$. The constants a and b are known and can be $+\infty$ or $-\infty$. And some of the methods are reviewed by Guenther(1978).

For the two-truncation parameter families, Bar-lev and Boukai(1985), Karkostas(1985) considered MVU estimation of an arbitrary functions of parameter. Krishnamoorthy et al(1989), have studied the estimation in the type-II censored sample from a one-truncation parameter family. Using two samples for one-truncation parameter family the MVU estimator is dealt with by Rohatgi(1989); Selvavel(1989) has studied the problem of estimation for two one-truncation parameter families, when both the samples are type-II censored.

Patel and Bhatt(1990) have considered a problem of finding MVUE, when sample is (singly) type-II censored and taken from two-truncation parameter families. Ferentions (1990) has considered the same but for a (doubly) type-II censored samples. Some of authors have constructed the confidence intervals for $g(\theta)$ and MVUE of $P(Y > X)$ in the two sample cases, which has been of great interest to researchers in reliability theory and survival analysis. A reasonably complete account of these papers would span over several chapters of the present thesis.

In the following section we have given the chapterwise summary in brief.

1.2 CHAPTERWISE SUMMARY :

The present thesis consists of five chapters. The first two chapters deal with the introduction and some preliminary definitions and results like unbiased estimator, MVU estimator, concept of truncation and censoring. Methods of constructing MVUEs supplimented by suitable examples are discussed in chapter 2.

In chapter 3, methods of constructing MVUEs of U-estimable function $g(\theta)$ have been discussed, when sample is taken from one-truncation parameter families. The same is discussed for type-II censored sample. Confidence interval for $q(\theta)$ in a censoresed sample has been studied.

Chapter 4 is devoted to two one-truncation parameter families. Different methods have been discussed for completeand type-II censored samples. The MVUE of the probability (Y > X) is discussed. Confidence interval for $q(\theta)$ in a censored sample has been studied.

Methods of constucting MVUEs of U-estimable function $g(\theta)$, when sample is taken from the two truncation parameter families. The same is considered for type-II censored sample. Confidence interval for $q(\theta)$ for a censored sample etc are the subject matter of chapter 5.

In this dissertation, definitions, theorems, lemmas, examples when numbered are governed by the chapter and the section in which they appear and are numbered serially within a section. Thus for example 2.2.1 refers to the example number 1 in chapter 2 scetion 2, etc. Similarly equations are serially numbered in each section. Also $a = b \cdot c$ means that a is equal to b multiplied by c.