CHAPTER 0

INTRODUCTION

There is no unified theory of testing statistical hypothesis from which all tests of significance can be deduced as acceptable solutions. In many situations test criteria may have to be obtained from intuitive considerations. One such theory, contributed by Neyman and Pearson (1933), marked an important development because it unfolded the various complex problems in testing statistical hypothesis, and led to the construction of general theories in problems of discrimination (identification), sequential analysis etc.

Statistical decision theory was introduced by Abraham Wald (1939) as a generalization of the classic statistical theories of hypothesis testing and estimation. After Wald's work important contributions have been made by Girshick and Savage (1951) and Stein (1956). A detail introduction is given by Ferguson (1967) and Berger (1980) in their texts.

The raw material of a statistical investigation is a set of observations, these are the values taken on by random variables X whose distribution P_w is at least partly unknown (w is parameter $w \in \Omega$). Statistical inference is concerned with methods of using this observational material to obtain information concerning the distribution of X or the parameter w. Suppose that a choice has to be made between a number of alternative actions. The observations, by providing information about the distribution from which they came, also provide guidance as to the best decision. The problem is to determine a rule which, for each set of values of the observations, specifies what decision should be taken. Mathematically such a rule is a function ∂, which assigns to each possible value $\underset{\sim}{x}$ of the random variables a decision $d = \partial(\underset{\sim}{x})$, that is a function whose domain is the set of values of X and whose range is the set of possible actions.

In order to see how ∂ shoud be chosen, one must compare the consequences of using different rules. Suppose that the consequence of taking decision d when the distribution of X is P_w is loss, which can be expressed as a non-negative real number $L(w,d)$. The expectation $E[L(w,\partial(\underset{\sim}{X})]$ evaluated under the assumption that P_w is the true distribution of $\underset{\sim}{X}$. This expectation which depends on the decision rule ∂ and the distribution P_w, is called the risk function of ∂ and will be denoted by $R(w,\partial)$. This suggests that select a decision function which minimizes the resulting risk $R(w,\partial)$.

If no prior information regarding w is available one might consider the maximum of the risk function its most important feature. Of two risk functions the one with smaller maximum is then preferable, and the optimum procedure is that with minimax property of minimizing the maximum risk. A minimax solution is one that gives the greatest possible protection against large losses. Such a principle may sometimes be quite unreasonable. The basic difference between the philosophy of the Baysian and non Baysian is that the Baysian considers the parameter of the distribution as the random variable, whereas non-Baysian regrads it as a fixed point.

This dissertation represents an attempt to summurize in an integrated form some of the results in the field of Bayes procedures for testing of hypotheses. It has a logical basis of its own and has an important part to play in drawing inferences from data.

As basic requirement to the study of Bayes test procedures Section 1.1 is devoted to a mathematical theory of hypothesis testing developed by Neyman and Pearson. Bayes procedures for testing of hypotheses is a statistical decision problem when parameter space and decision space contains only two points. Bayes test have a striking similarity with that of most powerful test (MP-test). In chapter I emphasis is given to illustrate the above statement and Bayesian likelihood criterion is established in Section 1.3. A mathematical result is given for giving Bayes test procedures in case of symmetric posterior density function and symmetric loss function. In Section 1.2 general decision problem is discussed together with Bayes risk, Bayes decision function based on the concept of prior distribution and posterior distribution is introduced. An attempt is made to illustrate the importance of taking observations in decision making, with the help of an example. When there are two decisions making a decision d_1 (making of decision d_2) is taken as accepting the hypothesis H_1 (H_2). For given prior distribution and decision function we can compute average risk. In order to obtain Bayes rule, minimize the same under two different hypotheses. Take a decision d_1 based on observations provided the posterior risk in taking decision d_1 is less, otherwise decision d_2 is taken. Some of the examples solved in Chpater I assumes the above. A comparative remark about classical and Bayes test procedure is also given.

Some useful definations used in chapter I are :

Defination : 0.1 :

Let $\{f_w, w \in \Omega\}$ be a family of pdf's (pmf's), $\Omega \subseteq R$. We say that f_w has a monotone likelihood ratio (MLR) in the statistic T(X) if for $w_1 < w_2$, whenever f_{w_1}, f_{w_2} are distinct, the ratio $\frac{f_{w2}(x)}{f_{w_1}(x)}$ is a non-decreasing function of T(X) for the set of values X for which at least one of f_{w_1} and f_{w_2} is > 0.

Defination : 0.2 :

If there exist real-valued- functions $Q(w)$ and $D(w)$ on Ω and Borel-measurable function $T(X_1, X_2, \ldots, X_n)$ and $S(X_1, X_2, \ldots, X_n)$ on R_n such that,

$$f_w(x_1, x_2, \ldots, x_n) = \exp\left\{ Q(w)T(\underset{\sim}{x}) + D(w) + S(\underset{\sim}{x}) \right\}$$

we say that the family $\{f_w, w \in \Omega\}$ is a one parameter exponential family.

Chapter II deals with Bayes test procedures for vector valued parameter and multiple hypotheses tests. In this chapter observations are considered in the form of vector. The loss associated with decision rule is defined with the help of norms. Different models are used . the first . considered is bivariate normal distribution with simple zero-one loss function and acceptance region for Bayes decision rule is sketched. Thereafter Bayes test procedures are obtained for various hypotheses and loss functions. In one of such models (2.1.1) risk function follows non-central χ^2 distribution and Bayes test procedure depends on non-centrallity parameter. In section 2.1.2 a k-dimensional multinominal random vector is considered and Bayes test procedure is obtained.

Section 2.2, is devoted to multiple hypotheses testing. In this section the parameter space is partitioned into m subsets. The decision that the parameter belongs to a particular subset of the parameter space is interpreted as the acceptance of that corresponding hypothesis, and rejection of the other m-1 alternative hypotheses. Also the procedure is illustrated with an example.

Chapter III deals with Bayes sequential test procedure. In Section 3.1 sequential sampling and related components have been introduced. It is also shown that there is 'gain' due to sequential sampling procedure as compared to that of fixed sample size procedure. Section 3.2 represents further ideas and technique of solving Bayes sequential testing problem. Also it is shown with the help of an example that Bayes sequential procedure need not always exist. In order to decide when to stop sampling in Bayesian sequential analysis, one has to compare the posterior Bayes risk of an immediate decision with the expected Bayes risk of continuing sampling optimally. The technique of Back-ward induction is used to find Bayes sequential decision procedure. An example to illustrate the procedure is also given.

Section 3.3 contains Wald sequential probability ratio test(SPRT) for two simple hypothesis. The intention to study the same is, to show that SPRT is Bayes procedure, which in-turn implies, the problem in which both the parameter space and decision space have exactly two points, the optimal sequential decision procedure is either to choose a decision immediately without any observations or to use SPRT.