Chapter-3.

BAYESIAN INFERANCE IN RELIABILITY (ADVANTAGE OVER CLASSICAL INFERANCE).

Bayes theorem is a foundation of Bayesian inferance. Using Bayes theorem, data points and preassumed information (prior information) regarding to parameter are linked together and is used for related inferance about parameter(s). In the recent times this method is becoming very popular in almost all areas of statistical applications.

The extent to which gains have been recognized are pointed out by Box and Tiao (1973), and for ready referance we quote, from page no. 2 of this book. ..."Bayesian inferance alone seems to offer the possibility of sufficient flexibility to allow reaction to scienyofic complixity free from impediment from purely technical limitation". Kendall and Staurt (1961) comments "The principle argument in favour of confidance intervals, however, is that they can be derived in terms of frequency theory of probability without any assumptions concerning prior distributions such as are essential to the Bayes approach. This, in our opinian, is undeniable. But it is fair to ask wheather they achive this economy of basic assumptions without losing something which the Bayes theory possesses. Our view is that they do, in fact, lose something on occasion, and that this something may be important for the purpose of estimation....". The effect of this loss in reliability estimation is sharp, to be illustrated with some examples in this chapter.

3.1 Foundations Of Bayesian Statistical Inferance :

I> Subjective Probability (Personal Probability) :

Frequency notion of probability deals with events and series of experiments (trials). In this situation these experements have to be conducted, under the same environment, as many times as we want. Probability of an event is the ratio of number of times that event occured to that of the total number of trials. From axiomatic definition of probability one can compute the probability of an event of interest. For example, to examine the unbisedness of a coin or for estimating the probability of head, one can conduct an experiment of tossing a coin and use the ratio of number of heads to total number of repetations.

While, subjective probability not only deals with events but also with some problems to be solved. Moreover the nature of the events is such that, it is not possible to repeat the experiment to compute the proportion. For example, an event like "nuclear power plant X will suffer a core meltdown". As a model one can pre-suppose the probability of such event and from statistical point of view, these are called hypothesis. Our interest is to test these hypothesis with certain degree of belief in it. As our evidance increases, relevent/against hypothesis, we change our degree of belief in hypothesis Once degree of belief about a proposition 'A' can be quantified by P(A); we speak in our daily life : "probabily I will get distinction", "He will probably marry her because...", etc. Two extremities of beliefs in proposition (event) 'A' are, it is true (that is, P(A) = 1) and it is false (that is, P(A) = 0). Intermediate beliefs are points from (0,1), for the values of P(A).

Once belief in a perticular hypothesis may be different from that of others. That is, the probability assigned by one individual may be quiet different from other individuals. Therefore, subjective probability some times refered to as "*personal probability*". As an investigator is rarely certain about the true nature of the process that generates observed events, assumptions are made about underlying process. Now, question arises about validity of these assumptions. For example, many times in reliability analysis we are assuming that data are generated according to Poisson process, even when there is little or almost no evidance to support such an assumption.

Subjective notion of probability is incarporated in Bayesian analysis by the prior distribution. It is the distribution of degree of belief about θ before data $\underset{\sim}{Y}$ is observed. As a result, prior distribution not admit a direct limiting frequency interpritation. However, in some situations, observed data may be used to estimate prior distribution. Past data sometimes used for prior estimation. For example : Proportion of defectives (say θ) in a sample are plotted against sample number so as to obtain prior distribution of θ. Wheather prior distribution does exists which is responsible for quantification in Bayesian analysis depends on nature of problem. It is well suited for use in reliability analysis.

Hence we conclude that subjectivity enters into all statistical analysis and that such analysis is an art as well the science. One of the branch which utilizes the subjectivity is Bayesian inferance.

II>Sampling Theory Verses Bayesian Inferance:

To demonstrate differance between sampling theory and Bayesian methods of inferance, consider an example : Suppose we want to study life length of a certain population of energy converter elements under spetial use conditions. Assume tentatively that observed lives of these elements are distributed independantly and exponentially with mean life θ. Let $(Y_1, Y_2, \ldots, Y_n)$ be a given data of size n from this population. Joint p.d.f. is given by

$$f(\underset{\sim}{y}/\theta) = (1/\theta)^n.\exp[-\sum_{i=1}^{n} y_i/\theta\ ] \qquad \ldots(3.1).$$

We are interested in making inferance about θ based on n data values.

3.1.1 Inferance Based On Classical Theory :

From classical theory (sampling theory) point of view, mean life θ is assumed to be fixed constant. Its estimator $\hat{\theta}(\underset{\sim}{Y})$, function of $\underset{\sim}{Y}$, can be obtained according to method of moments, maximum likelihood, minimum variance ir least squares, etc. For this example, here maximum likelihood and method of moment estimators of θ are same, is given by

$$\hat{\theta} = \hat{\theta}(\underset{\sim}{Y}) = \sum_{i=1}^{n} Y_i/n \qquad \ldots(3.2).$$

Imagining $\{Y_1, Y_2, \ldots\}$ set of all data points from hypothetical population (2.1), sampling distribution of $\hat{\theta}(\underset{\sim}{Y})$ is such that $U = 2n\hat{\theta}(\underset{\sim}{Y})/\theta$ has $\psi^2(2n)$, that is the p.d.f. of U is given by

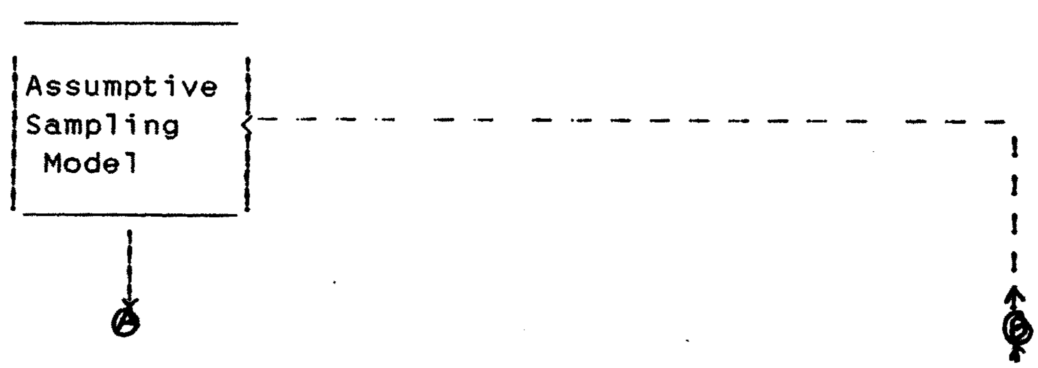
$$f(u) = (2^n \Gamma n)^{-1} . \exp\{-u/2\} . u^{(n-1)};\ 0 < u < \infty \ldots(2.3).$$

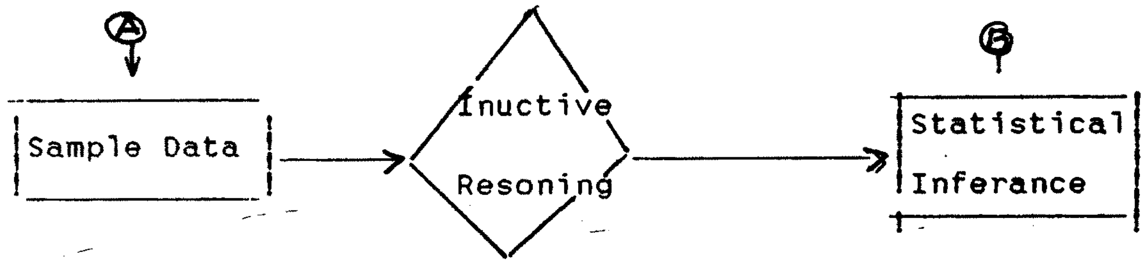
Two sided confidance interval (TCI) estimator for θ with confidance $(1-\gamma)\%$ is given by

$$[2n\hat{\theta}(\underset{\sim}{Y})/\psi^2_{(1-\gamma/2)}(2n);\ 2n\hat{\theta}(\underset{\sim}{Y})/\psi^2_{(\gamma/2)}(2n)] \quad \ldots(3.4).$$

Note that, in repeated sampling in which such a confidance interval is computed from each sample, the computed confidance intervals would include the true value θ in $(1-\gamma)\%$ of the cases. Confidance interval is not probability statement about θ, since θ is not a r.v. That is the probability that θ exceeds the upper limit of TCI is $\gamma/2$. In short, sampling theory inferances are of inductive reasoning, since from sample we are making inference about population.

For ready referance, following figer (from Martz and Waller (1982) on page no.166), gives brief picture about sampling inferance.





Inferance Based On Classical Theory.

3.1.2 Bayesian Inferance :

Compared to classical theory, Bayesian method is much more direct. Here instead of θ as a fixed constant (as in classical theory), it is assumed to be a non-observable random variable (r.v.) with prior p.d.f. $g(\theta)$. This prior density expresses the state of knowledge or ignorance about θ before sample data. Since very rarely parameters of interest are known, it is helpful to consider parameter as a r.v. and modeling for its distribution (uncertanity) is again an art as well the science.

For above example, assume that prior distribution of Θ is uniform on the range θ_1 to θ_2, whwre $0<\theta_1<\theta<\theta_2<\infty$. That is,

$$g(\theta) = (\theta_2-\theta_1)^{-1}; \quad 0<\theta_1<\theta<\theta_2<\infty. \qquad \ldots(3.5).$$

Let, $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_n)$ be a sample of size n from a population with p.d.f. $f(x/\theta)$. As the distribution of a $\underset{\sim}{Y}$ depends on θ; $\underset{\sim}{Y}$ does contain '*some*' information about the unknown parameter θ. Now by using information about θ gather through $\underset{\sim}{Y}$ the prior knowledge can up-dated, and this up-dated knowledge can be expressed by the posterior distribution of θ (which can be obtained by using Bayes theorem). For model described in (3.1) and (3.5) posterior distribution is given by

$$g(\theta/\underset{\sim}{y}) = \frac{(\Sigma y_i)^{n-1} \exp[-\Sigma y_i/\theta]}{\theta^n [\Gamma(n-1, \Sigma y_i/\theta_1) - \Gamma(n-1, \Sigma y_i/\theta_2)]} \quad \ldots(3.6).$$

where, $\Gamma(a,z) = \int_0^z y^{a-1} . \exp[-y].dy$, $a > 0$; and Σ stands for sumation over $i = 1,2,\ldots,n$.

Now, under squared error loss function point estimator of θ is mean of posterior distribution and is given by

$$E(\Theta/\underset{\sim}{y}) = \Sigma y_i \left[\frac{\Gamma(n-2, \Sigma y_i/\theta_1) - \Gamma(n-2, \Sigma y_i/\theta_2)}{\Gamma(n-1, \Sigma y_i/\theta_1) - \Gamma(n-1, \Sigma y_i/\theta_2)} \right]; \; n > 2. \quad \ldots(3.7).$$

Bayes interval estimator for θ is also obtained from posterior distribution of θ.

Definition 3.2.1 :

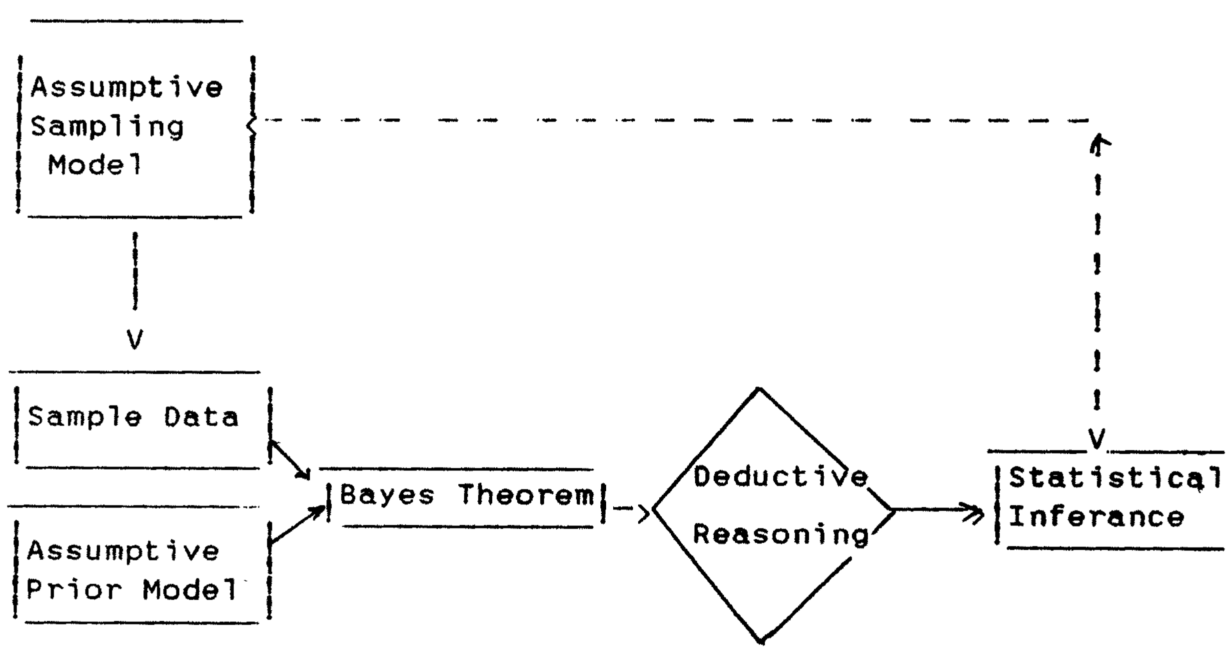
(θ_*, θ^*) is called $(1-\gamma)\%$ two-sided Bayes probability interval for Θ if,

$$\int_{-\infty}^{\theta_*} g(\theta/\underset{\sim}{y}).d\theta = \gamma/2 = \int_{\theta^*}^{\infty} g(\theta/\underset{\sim}{y}).d\theta \, . \quad \ldots(3.8).$$

Probability interval provided by above definition need not be shortest length interval, therefore one may go for that type of intervals. Shortly we will discuss some other kinds of probability intervals in last section of this chapter.

Differance between Bayesian analysis and classical theory approach is that, Bayesian analysis takes explicit account of prior distribution while in classical theory it is not. Sometimes Bayes and classical theory approaches may give same results but, there is a differance in the interpritation of the results. For example : confidance interval is not a probability statement about θ, while Bayesian interval is.

Following figer (from Martz and Waller (1982), p.p.168) gives brief idia of of Bayesian method of inferance :



Bayesian Method Of Inferance.

Two more distinctive features of Bayesian inferance are-

i> Due to exclusive use of sample data, classical theory inferance are more ristrictive than Bayes inferance. Whereas, Bayes inferance technique uses past information in terms quantified prior distribution, gives more informative inferance. Degree of more informative inferance depends upon quality of assessements incorporated in the prior distribution.

ii> Classical theory inferance requirs comparatively more set of data points to attain same quality of inferance as that of Bayesian inferance. That is why, in more expensive situations and cases where it is not possible to obtain more data, Bayes inferance is having its importance. These situations mostly occur in reliability estimation.

3.3 Bayesian Inferance In Reliability :

1 Some Drawbacks With Classical Theory In Reliability :

Sampling theory is found to be very important in reliability estimation, but many times it is found to be below our satisfaction. These are mentioned here.

Cost effectiveness reliability estimation is one of the major factor, which contradicts to use of classical theory. Due to limited time and funds available to run an experiment less no. of sample observations available. As a result low level of confidance in the reliability estimate. Growhaski, Hausman and Lamberson (1976) illustrated this with following example :
Suppose that in experiment, to determine reliability of a redesigned automobile airconditioning system only limited time and funds are available for testing single vehicle for 36,000 miles. If this experiment resulted into two failures then based on exponential model, a 90% two sided confidance interval estimate for the system reliability at 12,000 miles (say waranty ends at this period) is found to be (12.2%, 88.8%). (This confidance interval is due to, under exponential life time

model number of failures follows poisson process. Bain (1978) (Theorem (1.4.3),p.p.106) and Martz and Waller (1982) (Table 4.4, p.p.122)). Such an estimation is practically useless.

In similar manner classical estimation techniques are useless in caes of scarce data. In case of highly reliable system, failure data will be scarce. For example : Data from Nuclear Regularitory Commission (1978) reported zero value failures over $7.9*10^{6}$ hours for population of 16 boiling water reactors. As a result under exponential failure time model, point estimate of constant failure rate is zero, highly optimistic result and confidance interval here is not possible to obtain. Also frequency of core meltdowns in nuclear power plants are very rare over long period. Hence the same kind of problem arises as that in the above example for estimation of failure reactors of nuclear power plant due to core meltdowns.

Most of the engineering designs are evolutionary rather than revolutionary. That is, engineering equipments get modified in its old design so as to fulfil new requirements from customers. This modified design must be reliable at least as that of the old one. In classical theory it is not possible to incorporate such previous information. As an exmple : Suppose previous year model equipment yielded an observed system reliability of 85% based on warranty data. The fact, new design is an evolution of old design, which had an observed reliability of 85%, is an important consideration that can not be taken into account using methods based on classical theory.

Sometimes classical theory methods gives absurd results. If parameter of interest is known to lie within a specified range it is difficult to consider this information for classical methods. For example : (Martz and Waller (1982) p.p.171), suppose we are interested in estimating in meanlife (θ) of a certain automobile system, such as steering, brakes etc.based on the assumption of an exponential model. Generally, it is known that mean life of such automative systems exhibit a mean life between 10^4 and 10^5 miles per failure. Based on sample test data, classical theory can do no more than compute a confidance interval estimate according to (3.4). Such estimators are still true in the required proportion of cases, but the statements take no account of our prior knowledge about the range of θ and may occasionally be idle. It may be true, but would be absured to say that 10^3 to 10^6 is confidance interval even we know that θ lies between 10^4 to 10^5.

Sometimes due to classical theory, overdesigning and high margines of protection are given than really it is needed. Ultimately, again it results in to large no. of tests and long duration of time. Higher consumers and producers risks have been highlighted. As an example : Suppose the mean life (θ)of an exponential distribution is a random variable and reliability demonstration test is used with θ_1 (minimum acceptable MTTF) and δ = 10% ;where, δ = P(accepting H_0 under H_1 is true) :consumers risk; but, suppose it is known that $P(\Theta \leq \theta_1)$ =

0.0001. In this case one is paying (in terms of sample size) to provide 10% protectionfor an event that is quiet rare.

II Advantages Of Bayesian Inferance In Reliability :

All statistical inferance theories, weather classical theory, Bayesian, likelihood or otherwise, require some degree of subjectivity in there use. For example : Classical theory analysis of (3.1) proceeds by assuming priori that data were *exactly* exponentially distributed with unknown parameter λ, that each observation had *exactly* same mean life and that each observation was distributed *exactly* independently of every other sample observation. The Bayes method provides flexible and satisfactory way by assuming prior knowledge or ignorance, These assumptions lead via Bayes theorem to posterior inferances, that is, inferance obtained by incorporating data into analysis, about reliability parameter(s) of interest. Bayes theorem provides simple and error-free truism for incorporating prior information. The engineers generally appriciates to tell such prior information in formulized way.

Since engineering situations are evolutionary rather than revolutionary, subjectivity is basic to reliability engineering. This fact is personnel, so quantification of subjectivity may not agree with that of other engineer(s). One may worry about the fact of different prior assumption results into different inferances about the same problem. However, truthfullness in the subjective belief is an advantage of Bayesian inferance. If the group of engineers hold the same degree of belief then such

12651

argument reassure the reliability analyst that the resulting inferances are probabily correct. On the other hand, situatins where disagreements amoung the engineers, reliability analyst, in this case, either to ignore the judgement of one or more of the engineers, or further data be obtained to solve the problem. In any case, Bayesian inferance shows to what extent different results may or may not be obtained according to differances in prior options held. This credits Bayesian reliability analysis.

Two important practical benifits of Bayesian analysis are-

i> Under the condition that prior information accurately shows the true variation in the parameter(s), Bayesian analysis gives high quality of inferance, and

ii>Reduction in testing requirment, namely test time and sample size.

There is another important advantage of Bayesian inferance,namely, unacceptable inferances must come due to uncorrect assumptions and not from weakness (insufficiency) of method used to provide the inferances. From this point of view, since Bayesian inferance is more direct, it shows many drawbacks of sampling theory. Based on component data, system reliability analysis is possible in Bayesian inferance, since it is possible to manipulate probability statements on components into corrosponding statements on system reliability.

For previous automobile airconditioning example, provided by Grohowski, Hausman and Lamberson (1976), they constructed a

Bayesian interval estimate for system reliability at timt 12,000 miles to be (66%, 84%) with a point estimate of 75%. Compared to the previous model observed value of 78%, the Bayesian result, indicate that the system on the average is not quiet as good as the old, but is a fogur that is believable. On the other hand, the classical estimate of 51% is so low as to be absurdly pessimestic. Thus, Bayesian method produces believable results convincing engineers.

3.4 Nature Of Bayesian Inference :

3.4.1. Bayes Theorem :

Suppose that $\underset{\sim}{y} = (y_1, y_2, \ldots, y_n)$ is a vector of n observations whose probability distribution $p(\underset{\sim}{y}/\theta)$ depends on the values of k parameters $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$. Suppose also that $\underset{\sim}{\theta}$ itself has a probability distribution $p(\underset{\sim}{\theta})$. Then,

$$p(\underset{\sim}{y}/\underset{\sim}{\theta}).p(\underset{\sim}{\theta}) = p(y,\underset{\sim}{\theta}) = p(\underset{\sim}{\theta}/\underset{\sim}{y}).p(\underset{\sim}{y}) \qquad \ldots(3.9).$$

Given the observed data $\underset{\sim}{y}$, conditional distribution of θ is given by

$$p(\underset{\sim}{\theta}/\underset{\sim}{y}) = \frac{p(\underset{\sim}{y}/\underset{\sim}{\theta}).p(\underset{\sim}{\theta})}{p(\underset{\sim}{y})} \qquad \ldots(3.10).$$

Also we can write,

$$p(\underset{\sim}{y}) = \begin{cases} \int_{R_{\underset{\sim}{\theta}}} p(\underset{\sim}{y},\underset{\sim}{\theta}).d\underset{\sim}{\theta} & \text{,if } \underset{\sim}{\theta} \text{ continious;} \\ \sum_{\underset{\sim}{\theta}} p(\underset{\sim}{y},\underset{\sim}{\theta}) & \text{,if } \underset{\sim}{\theta} \text{ discrete.} \end{cases}$$

$$= \begin{cases} \int_{R_{\underset{\sim}{\theta}}} p(\underset{\sim}{y}/\underset{\sim}{\theta}).p(\underset{\sim}{\theta}).d\underset{\sim}{\theta} , & \text{if } \underset{\sim}{\theta} \text{ continious;} \\ \sum_{\underset{\sim}{\theta}} p(\underset{\sim}{y}/\underset{\sim}{\theta}).p(\underset{\sim}{\theta}) , & \text{if } \underset{\sim}{\theta} \text{ discrete.} \end{cases} \quad ...(3.11).$$

The statement (2.10) is rereferred to as Bayes theorem. In this expression, $p(\underset{\sim}{\theta})$, which tells us what is known about $\underset{\sim}{\theta}$ without knowledge of the data, is called the prior distribution of $\underset{\sim}{\theta}$, correspondingly, $p(\underset{\sim}{\theta}/\underset{\sim}{y})$, which tells us what is known about $\underset{\sim}{\theta}$ given knowledge of the data, is called the posterior distribution of $\underset{\sim}{\theta}$ given $\underset{\sim}{y}$. the quantity $p(\underset{\sim}{y})$ is "normalizing" constant necessary to ensure that the posterior distribution $p(\underset{\sim}{\theta}/\underset{\sim}{y})$ integrates or sum to one.

3.4.2 Bayes Theorem And The Likelihood Function:

Given the data, $\underset{\sim}{y}$, $p(\underset{\sim}{y}/\underset{\sim}{\theta})$ in (3.10) may be regarded as a function of $\underset{\sim}{\theta}$ and not of $\underset{\sim}{y}$, when so regarded, following Fisher (1922), it may called the likelihood function of $\underset{\sim}{\theta}$ for given $\underset{\sim}{y}$ and can be written as $l(\underset{\sim}{\theta}/\underset{\sim}{y})$. Bayes formula can be written as

$$p(\underset{\sim}{\theta}/\underset{\sim}{y}) = c.l(\underset{\sim}{\theta}/\underset{\sim}{y}).p(\underset{\sim}{\theta}) \quad ...(3.12).$$

where, $c = 1/p(\underset{\sim}{y})$: normalizing constant.

In other words, Bayes theorem tells us that the probability distribution for $\underset{\sim}{\theta}$ posterior to the data $\underset{\sim}{y}$ is proportional to the product of the distribution for $\underset{\sim}{\theta}$ prior to the data and the liklihood for $\underset{\sim}{\theta}$ given $\underset{\sim}{y}$, that is,

Posterior Distribution α Likelihood * Prior Distribution.

The likelihood function $l(\underset{\sim}{\theta}/\underset{\sim}{y})$ playes a very important role in the Bayes formula. It is the function through which the data $\underset{\sim}{y}$ modifies the prior knowledge of $\underset{\sim}{\theta}$. Therefore it can be thought of as representing the information about $\underset{\sim}{\theta}$ comming from the data.

The likelihood function is defined up to a multiplication constant,that is, multiplication by a constant leaves the likelihood unchanged. This is in accord with the role it playes in Bayes formula; since multiplying the likelihood function by an arbitrary constant will have no effect on the posterior distribution of $\underset{\sim}{\theta}$. The constant will cancel upon normalizing the product on the right hand side of (3.12). This relative value of likelihood is of importance.

3.4.3 The Standardised Likelihood :

When the integral $\int l(\underset{\sim}{\theta}/\underset{\sim}{y}).d\underset{\sim}{\theta}$, taken over the admissible range of $\underset{\sim}{\theta}$, is finite, then occassionally it will be convinient to refer the quantity

$$\frac{l(\underset{\sim}{\theta}/\underset{\sim}{y})}{\int l(\underset{\sim}{\theta}/\underset{\sim}{y}).d\underset{\sim}{\theta}} \qquad \ldots(3.13).$$

We shall call this the standerdised likelihood, that is, the likelihood scaled so that the area, volume, hypervolume under the curve, surface, or hypersurface, is one.

3.4.4 Sequenyial Nature Of Bayes Theorem :

Note that, (2.12) provides a mathematical formulation of how previous knowledge be combined with new knowledge (knowledge after data obtained). Also this allows us to update information about $\underset{\sim}{\theta}$ continually, as more observations are taken.

Thus, suppose we have an initial sample of observations $\underset{\sim}{y}_1$ then Bayes formula gives

$$p(\underset{\sim}{\theta}/\underset{\sim}{y}_1) \propto .p(\underset{\sim}{\theta}).l(\underset{\sim}{\theta}/\underset{\sim}{y}_1) \qquad \ldots(3.14).$$

Now, suppose we have a second sample of observations $\underset{\sim}{y}_2$ distributed independently of the first sample, then

$$p(\underset{\sim}{\theta}/\underset{\sim}{y}_1,\underset{\sim}{y}_2) \propto p(\underset{\sim}{\theta}).l(\underset{\sim}{\theta}/\underset{\sim}{y}_1).l(\underset{\sim}{\theta}/\underset{\sim}{y}_2)$$

$$\propto p(\underset{\sim}{\theta}/\underset{\sim}{y}_1).l(\underset{\sim}{\theta}/\underset{\sim}{y}_2) \qquad \ldots(3.15).$$

Observe, (3.15) is precisely of the same form as (3.14) except that $p(\underset{\sim}{\theta}/\underset{\sim}{y}_1)$, the prior distribution for $\underset{\sim}{\theta}$ given $\underset{\sim}{y}_1$, playes the role of prior distribution for the second sample. Obviously, this process can be repeated any number of times. In perticular, if we have n independant observations, posterior distribution can be recalculated after each new observation, if we want. Therefore, at the m^{th} stage the likelihood associated with m^{th} observation is combined with the posterior distribution of $\underset{\sim}{\theta}$ after (m-1) observations to give the new posterior which is now based on m observations, given by

$$p(\underset{\sim}{\theta}/\underset{\sim}{y}_1,\underset{\sim}{y}_2,\ldots,y_m) \propto p(\underset{\sim}{\theta}/\underset{\sim}{y}_1,y_2,\ldots,y_{m-1}).l(\underset{\sim}{\theta}/\underset{\sim}{y}_m);\ m = 2,3,\ldots \quad \ldots(3.16).$$

where, $p(\underset{\sim}{\theta}/\underset{\sim}{y}_1) \propto p(\underset{\sim}{\theta}).l(\underset{\sim}{\theta}/\underset{\sim}{y}_1)$.

Thus, in fundamental way, Bayes theorem describes the process of learning from experiance and shows how knowledge about the state of nature represented by $\underset{\sim}{\theta}$ is continually modified as new data becomes available.

Following example, (from Martz and Waller (1982)), dipicts how likelihood improves the prior judgements made by two engineers A and B about their ability to achive mean life (θ) of engine in different intervals.

Suppose two design engineers A and B are given the task of redesigning an industrial engine that is to have a mean life θ of at least 3000 hour. Engineer A has had considerable experiance in the design of similar engines and can make a moderately good guess of the success of the effort. On the other hand, B has had less experiance and is far less certain of the outcome of the task. With the help of a reliability analyst, both A and B have been seperately encouraged to quantify their degree of belief in the success of their task. A and B quantified their beliefs in the following manner.

TABLE : a

θ (hour)	$g_A(\theta)$	$g_B(\theta)$
0 - 1000	0.01	0.15
1000 - 2000	0.04	0.15
2000 - 3000	0.20	0.20
3000 - 4000	0.50	0.20
4000 - 5000	0.15	0.15
> 5000	0.10	0.15

It is observed that A believes a priori that the probability is 0.75 that the design will be successful, while B believes that the probability is 0.50. Engineer A further believes that there is only a priori 0.05 probability that the design effort will grossely fail (a mean life less than 2000 hour.), whereas B believes that the prior probability that this will happen is 0.30. In other words, the past experiance of A leads him to a more optimestic view concerning the success of the task, whereas B has a more pessimistic attitude.

Following the redesign effort the reliability analyst propose that two prototype engines be tested untill both fail. Assuming an exponential failure time distribution with mean life (θ), the standerdised likelihood of θ is given by

$$l(\theta/x_1,x_2) = \frac{f(x_1,x_2/\theta)}{\int f(x_1,x_2/\theta).d\theta}$$

$$= (x_1+x_2)/\theta^2.\exp\{-(x_1+x_2)/\theta\} \quad ...(3.17).$$

$$0< \theta< \infty.$$

Now suppose that the test has been conducted and that $x_1 = 2000$, and $x_2 = 2500$ are observed. The standerdized likelihood that θ is between a and b, where $a< \theta< b$, is given by

$$l(a< \theta<b/2000,2500) = \int_a^b (4500)/\theta^2.\exp\{-(4500)/\theta\}.d\theta$$

$$= \exp\{-(4500)/b\} - \exp\{-(4500)/a\},$$

$$...(2.18).$$

Using this we can easily compute the values of the standerdized likelihood to be as follows.

TABLE : b.

θ (hour).	$l(\theta/2000,2500)$.
0 -1000	0.01111
1000 - 2000	0.10
2000 - 3000	0.12
3000 - 4000	0.10
4000 - 5000	0.08
> 5000	0.59

Note that the likelihood is 77% that θ exceeds requirment 3000 hour, in spite of the fact that neither engine survived 3000 hour. This is due to the exponentiality and associated property which says that, in such a model, 63% of the failures are expected to occure prior to the mean life θ. Consequently, it is expected that the mean life should exceed 5000 hour, as indicated by the likelihood of 59%, given x_1 = 2000 and x_2 = 2500.

Now let us compute the posterior distribution for A in light of test results. According to Bayes theorem

$$g_A(0< \theta< 1000/2000,2500) = \frac{l(0 <\theta <1000).g_A(0< \theta< 1000)}{f_A(2000,2500)}.$$

where, $f_A(2000,2500)$ denotes the standerdized marginal distribution for A, evaluated at x_1 = 2000 and x_2 = 2500. It may be calculated according to

$$f_A(2000,2500) = \int_0^{\infty} (4500)/\theta^2.\exp[-(4500)/\theta].g(\theta).d\theta$$

$$= [\int_0^{1000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.01) +$$

$$[\int_{1000}^{2000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.04) +$$

$$[\int_{2000}^{3000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.20) +$$

$$[\int_{3000}^{4000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.50) +$$

$$[\int_{4000}^{5000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.15) +$$

$$[\int_{5000}^{\infty} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.10)$$

$$= (0.01).\exp[-4500/1000] +$$

$$(0.04).[\exp[-4500/2000] \exp[-4500/1000]]+$$

$$+..... +(0.01).[1-[\exp[-4500/5000]].$$

$= 0.15.$

Thus,

$$g_A(0< \theta< 1000/2000,2500) = (0.01111).(0.01)/(0.15)$$
$$= 7.4*10^{-4}.$$

Similarly, by computing remaining posterior probabilities for A we tabulate them as follows.

TABLE : c

θ (hour)	$g_A(\theta/2000,2500)$
0 - 1000	$7.4*10^{-4}$
1000 - 2000	$2.5*10^{-2}$
2000 - 3000	0.16
3000 - 4000	0.34
4000 - 5000	0.08
> 5000	0.40

It is noted that, in light of the observed data, the posterior distribution of A indicates that the engineer now belives that there is 0.82 probability that θ exceeds the requirment, upward from 0.75 prior to the data. By comparing table values in *TABLE : a* and *TABLE : c*, it is observed that A'^{s} strong belief that $3000< \theta<4000$ has been reduced from 0.50 to 0.34, whereas the bekief that $\theta > 5000$ has increased in view of the test results from 0.10 to 0.40. We see that in this case, neither the prior nor the likelihood "dominates" the posterior, but that both are fairly equally weighted in the analysis. This is not always the case as will now be shown.

The posterior distribution for B is similerly computed by use of Bayes theorem. For example

$$g_B(0< \theta< 1000/2000,2500) = \frac{l(0 <\theta <1000).g_B(0< \theta< 1000)}{f_B(2000,2500)}.$$

where, $f_B(2000,2500)$ denotes the standerdized marginal distribution for B, evaluated at x_1 = 2000 and x_2 = 2500. It may be calculated according to

$$f_B(2000,2500) = \int_0^{\infty} (4500)/\theta^2.\exp[-(4500)/\theta].g(\theta).d\theta$$

$$= [{}_0\int^{1000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.15) +$$

$$[{}_{1000}\int^{2000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.15) +$$

$$[{}_{2000}\int^{3000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.20) +$$

$$[{}_{3000}\int^{4000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.20) +$$

$$[{}_{4000}\int^{5000} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.15) +$$

$$[{}_{5000}\int^{\infty} (4500)/\theta^2.\exp[-(4500)/\theta]].(0.15)$$

$$= (0.15).\exp[-4500/1000] +$$

$$(0.15).[\exp[-4500/2000] \exp[-4500/1000]]+$$

$$..... +(0.15).[1-[\exp[-4500/5000]].$$

$$= 0.16.$$

Thus,

$$g_B(0< \theta< 1000/2000,2500) = (0.15).(0.011111)/(0.16)$$

$= 0.01.$

Similarly, by computing remaining posterior probabilities for B we tabulate them as follows.

TABLE : d

θ (hour)	$g_B(\theta/2000,2500)$
0 - 1000	0.01
1000 - 2000	0.09
2000 - 3000	0.15
3000 - 4000	0.13
4000 - 5000	0.07
> 5000	0.55

In view of the observed data, B now belives that there is 0.75 probability that $\theta > 3000$, as apposed to the prior probability of 0.50. Comparing *TABLE : a* and *TABLE : d* the likelihood is thus observed to have a fairly influential effect on B'^{s} posterior distribution. In fact, apart from a slight increase in the centre of the posterior distribution due to the slight influence of the prior, the posterior resembles the likelihood. In such a case we say that the prior is "dominated" by the likelihood.

In general, the sharpness or flatness of the prior distribution relative to the sharpness or flatness of the likelihood determines wheather the prior dominates the likelihood. Generally, if the prior is flat relative to the likelihood, then likelihood dominates the prior; as we see in the case of engineer A'^{s} prior in above example. On the other hand, if likelihood is flat relative to prior, prior dominates the likelihood. Box and Tiao (1973) point out that, generally in case of analysing scientific data, the likelihood dominates the prior and rarely apposite holds.

3.5 Performig A Bayesian Reliability Analysis :

Definition : Martz and Waller (1982),

" A Bayesian reliability analysis consists of the use of statistical methods in reliability problems that involve the parameter estimation in which one or more parameters are considered to be a r.v. with a nondegenerate prior probability distribution which expresses the analyst's prior degree of belief about the parameters".

3.5.1 Problems and Considerations In A Bayesian Reliability Analysis :

In performing Bayesian reliability analysis, identification, selection and justification of prior distribution is very important and is difficult too. Following problems arises while doing the Bayesian analysis :

i> Which prior distribution to use.

ii> What sources of data are available for selecting a prior model.

iii> How to quantify the subjective information, and

iv> Which procedure is appropriate for fitting prior distribution to subjective data.

When multiple sources of relevant data are available for analysis, it must be decided that which data are to be used in fitting the prior distribution and which data are to be used in the likelihood function. Even this is not an easy task, but traditionally softer and more objective data have been allocated to prior, whereas the harder and more objective sample data have been used in likelihood.

Identification of prior concerns with wheather to use discrete or continuous prior, informative or noninformative prior, its mathematical simplicity and convenience, flexiblity of family, degree of softness or hardness of the subjective data sources.

3.5.2 Preposterior Analysis :

Preposterior analysis is a procedure for analysing (or searching) a prior distribution, before achivement of test data. Based on its (tentative priors) impact (effect) on posterior with respect to hypothetical data, desirability of this tentative prior is to be checked (studied).

Followings are steps for preposterior analysis : -

i> Initially select a tentative prior distribution.

ii> For a set of hypothetical data obtain posterior distribution.

iii> Study the set of posterior distribution so that wheather they seem reasonable in light of hypothetical data.

iv> If they are reasonable, prior distribution (which are tentatively selected) is good, otherwise go for another prior and repeat the steps untill it is resonable.

3.6 Bayesian Decision Theory :

Statistical decision theory concerns with making decisions about state of nature θ, which is unknown, based on given set of data. Decidion maker has to make a choice from a given set of available actions $\{a_1, a_2, \ldots, a_k\}$. In Bayesian decision theory, Θ is assumed to have a prior distribution. The decision maker combines prior knowledge of Θ and information provoded by an experiment by posterior distribution of Θ. Then he chooses an action so as to minimize expected loss over posterior distribution.

Notations :

1. Ω_θ : $\{ \theta \}$: parameter space or possible states of nature (may be vector valued).
2. $\mathcal{A}$: $\{ a \}$: action space.
3. $L(\theta, a)$: loss due to action a when state of nature is θ.
4. $\mathcal{X} = \{ \underset{\sim}{x} \}$: sample space of r.v. X which has p.d.f. $f(x/\theta)$,

when θ is true state of nature.

5. $\mathcal{D} = \{ \delta(\underset{\sim}{x}) \}$: decision space of possible decisions functions defined on $\mathcal{X}$ to action space $\mathcal{A}$.

3.6.1 Risk :

Note that, our action a depends upon particular sample data $\underset{\sim}{x}$ that we observe. For decision function δ, loss function may be written as $L(\theta, \delta(\underset{\sim}{x}))$, when we observe $\underset{\sim}{X} = \underset{\sim}{x}$ as a sample data. Therefore, loss function is a r.v. which depends upon sample outcome. Thus, risk is defined as expected value of the loss function. Mathematically it is given by

$$R(\theta,\delta) = E\left\{L(\theta, \delta(\underset{\sim}{x}))/\theta\right\}$$

$$= \int_{\mathcal{X}} L(\theta, \delta(\underset{\sim}{x})).f(\underset{\sim}{x}/\theta).d\theta \quad \ldots(3.19).$$

Naturally, good decision function would be one that minimizes the risk for all values of θ in Ω_θ. This need not be always possible to obtain such a decision function.

Definition 3.2.2: Inadmissible Decision Function :

A decision function δ^* is inadmissible if there exists any δ in $\mathcal{D}$ such that if,

$R(\theta,\delta) \leq R(\theta,\delta^*)$, for all values of θ in Ω_θ

and $R(\theta,\delta) < R(\theta,\delta^*)$, for at least one value of θ in Ω_θ. ■

Definition 2.3 : Better Decision Function :

δ^* is said to be better decision function than any δ in $\mathcal{D}$ if

$R(\theta,\delta^*) < R(\theta,\delta)$, for all values of θ in Ω_θ. ■

Definition 3.2.4 : Best Decision Function :

A decision function δ^* is best amoung the class of all decision functions D if

$R(\theta,\delta^*) < R(\theta,\delta)$, for all values of δ in D , δ's different from δ^*, for that value of θ ■

3.6.2 Bayes Risk :

Bayes risk of a decision function δ is defined as the expected value of the risk $r(\theta, \delta)$, with respect to the prior distribution g on Ω_θ.

Mathematically, Bayes risk $r(g, \delta)$ of decision function δ with respect to prior g is given by

$$r(g, \delta) = E\left\{R(\Theta, \delta)\right\}$$

$$= \int_{\Omega_\theta}\int_{\mathcal{X}} L(\theta,\delta(\underset{\sim}{x})).f(\underset{\sim}{x}/\theta).g(\theta).d\underset{\sim}{x}.d\theta \quad ...(3.20).$$

Thus, Bayes risk orders the decision space $\mathcal{D}$, that is, decision maker preferes decision function δ_2 to δ_1 if δ_2 has smaller Bayes risk than that of δ_1.

3.6.3 Bayes Decision Function :

Since Bayes risk orders the decision space $\mathcal{D}$, according to this principle one may search for a decision function δ, with respect to prior g, so that it has minimum Bayes risk. If such a decision function, say $\delta_g(\underset{\sim}{x})$, exists it is known as Bayes decision function and its associated Bayes risk given by

$$r(g) = r(g,\delta_g) = \min_{\delta\in D}\left\{ r(g, \delta) \right\} \qquad \ldots(3.21).$$

is known as the minimun Bayes risk.

Construction Of Bayes Decision Function :

Assuming, order of integration can be reversed, Bayes risk becomes

$$r(g, \delta) = \int_{\mathcal{X}} f(\underset{\sim}{x}) \left\{ \int_{\Omega_\theta} L(\theta, \delta(\underset{\sim}{x})).g(\theta/\underset{\sim}{x}).d\theta \right\} d\underset{\sim}{x}. \qquad \ldots(3.22).$$

where, $g(\theta/\underset{\sim}{x})$ is posterior probability distribution of Θ given $\underset{\sim}{x}$.

Therefore, to minimize the Bayes risk, a decision function $\delta(\underset{\sim}{x})$ should be chosen such that the inner integral is minimum, which is equal to expectation of $[L(\Theta,\delta(\underset{\sim}{x})/\underset{\sim}{x}]$ under posterior distribution of Θ given $\underset{\sim}{x}$, or simply posterior risk. Thus with respect to prior g, Bayes decision function can be obtained without going for value of minimum Bayes risk.

Let, $\Phi(\delta,\underset{\sim}{x})$ be posterior risk and ,

$$\phi(\underset{\sim}{x}) = \phi(\delta, \underset{\sim}{x}) = \min_{\delta\in\mathcal{D}} \phi(\delta, \underset{\sim}{x})$$

$$= \min_{\delta\in\mathcal{D}} E\left\{L[\Theta,\delta(\underset{\sim}{x})/\underset{\sim}{x}]\right\} \qquad \ldots(3.23).$$

Therefore minimum Bayes risk is given by

$$r(g) = \int_{\mathcal{X}} \phi(\underset{\sim}{x}).f(\underset{\sim}{x}).d\underset{\sim}{x}.$$

$$= E[\phi(\underset{\sim}{X})] \qquad \ldots(3.24).$$

3.7 Bayesian Estimation Theory :

Estimation problem is a perticular case of statistical decision problem, in which decision made by analyst is the estimate of an unknown parameter θ.

3.7.1 *Bayesian Point Estimation :*

Bayesian point estimation consists of,

$\mathcal{A}$:action space - set of point "estimates" of the parameter θ, which is subset of parameter space Ω_{θ}.

$\mathcal{D}$: decision space - possible "estimators" for θ.

Thus, decision function is $\hat{\theta} = \delta(\underset{\sim}{x})$, ...(3.25).

where, $\underset{\sim}{x}$ is observed value of $\underset{\sim}{X}$. On observing $\underset{\sim}{X}$, the function $\delta(\underset{\sim}{x})$ can be evaluated and the resulting "action" $\hat{\theta}$ is called point estimate of θ. The function $\hat{\theta}$ is called a point estimator for θ. Thus "estimates" (the actions) are the values of an "estimator" (a decision function).

While estimating θ by $\hat{\theta}$, the differance between the values of θ and its estimate $\hat{\theta}$ causes a loss. Thus, we should define a loss function such that for $\hat{\theta} = \theta$, loss is zero. For a reason as such, loss function in estimation problem is generally assumed of the form

$$L(\theta,\hat{\theta}) = \omega(\theta).d(\theta-\hat{\theta}), \qquad ...(3.26).$$

Where, d is nonnegative function of discrepancy $(\theta-\hat{\theta})$, such that $d(0) = 0$, and ω is weighing function that shows relative seriousness of a given discrepancy for different θ.

When the parameter is one dimentional, the loss function in an estimation problem can often be expressed as

$$L(\theta, \hat{\theta}) = \omega \mid \theta-\hat{\theta} \mid^{d}, \qquad \ldots(3.27).$$

where, ω, $d > 0$. When $d = 2$, the loss function is quadratic and is called squared-error loss function. If $d = 1$, loss function is called absolute error loss function, which is proportional to the absolute value of estimation error.

3.7.1.1 Bayesian Estimation Under Squared-Error Loss Function :

For one dimentional parameter, the squared-error loss function is given by

$$L(\theta,\hat{\theta}) = a\ (\theta-\hat{\theta})^2. \qquad \ldots(3.28).$$

Thus, Bayes risk of this loss function is given by

$$EL(\theta,\hat{\theta}) = a.E_{X/\theta}\left\{(\theta-\hat{\theta})^2/\theta\right\}$$

As we discussed in section 2.2 that there need not always exists a decision function (here $\hat{\theta}$) which minimizes risk for all values of θ. So, we have to choose such a rule, $\hat{\theta}$, which minimizes the expected value of $EL(\theta,\hat{\theta})$ under prior distribution of θ.

Now, from (2.22), $\hat{\theta}$ should be chosen such that,

$$\int_{\Omega_\theta} a.(\theta-\hat{\theta})^2.g(\theta/\underset{\sim}{x}).d\theta,$$

is minimum.

Which can be rewritten as

$$E_{\Theta/\underset{\sim}{x}}\left[a(\Theta - \hat{\theta})^2/\underset{\sim}{x}\right] = E_{\Theta/\underset{\sim}{x}}\left[a(\{\Theta - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\} - \{\hat{\theta} - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\})^2/\underset{\sim}{x} \right]$$

$$= a\, E_{\Theta/\underset{\sim}{x}}\left[\{\Theta - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\}^2 / \underset{\sim}{x}\right]$$

$$+a.\ E_{\Theta/\underset{\sim}{x}}\left[\{\hat{\theta} - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\}^2 / \underset{\sim}{x} \right]$$

$$-2a.\ \{\hat{\theta} - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\}.E_{\Theta/\underset{\sim}{x}}\left[\{\Theta - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\} / \underset{\sim}{x} \right]$$

$$= a\, E_{\Theta/\underset{\sim}{x}}\left[\{\Theta - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\}^2 / \underset{\sim}{x}\right]$$

$$+a.\ E_{\Theta/\underset{\sim}{x}}\left[\{\hat{\theta} - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\}^2 / \underset{\sim}{x} \right]$$

$$= a.Var_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x}) + a.E_{\Theta/\underset{\sim}{x}}\left[\{\hat{\theta} - E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})\}^2 / \underset{\sim}{x} \right] \qquad \ldots(3.29).$$

which will be minimum for $\hat{\theta} = E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})$

$$= \int_{\Omega_\theta} \theta.g(\theta/\underset{\sim}{x}).d\theta; \qquad \ldots(3.30).$$

which is mean of the posterior distribution.

Thus under squared-error loss function, Bayes estimator is simply the posterior mean of Θ given $\underset{\sim}{x}$. The minimum posterior risk associated with this Bayes estimator is equal to $a.Var_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})$. Therefore, according to (3.23) and (3.29),

$$\phi(x) = \min_{\hat{\theta}} E_{\Theta/\underset{\sim}{x}}\left[a(\Theta - \hat{\theta})^2 / \underset{\sim}{x}\right]$$

$$= a.Var_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x}). \qquad \ldots(3.31).$$

By (3.24), minimum Bayes risk associated with $\hat{\theta} = E_{\Theta/\underset{\sim}{x}}(\Theta/\underset{\sim}{x})$ is given by

$$r(g) = a.E_{\underset{\sim}{\Theta}/x}\left[\mathrm{Var}_{\underset{\sim}{\Theta}/x}(\Theta/\underset{\sim}{x})\right], \qquad \ldots(3.32).$$

which is Bayes risk of Bayes estimator.

3.7.1.2 Bayesian Estimation Under Quadratic Loss Function :

When parameter to be estimated is k dimentional vector, say $\underset{\sim}{\theta} = (\theta_1,\theta_2,\ldots,\theta_k)$, $k\geq 2$. A generalisation of squared error-loss function is the quadretic loss function which is given by

$$L(\underset{\sim}{\theta},\underset{\sim}{\hat{\theta}}) = (\underset{\sim}{\theta}-\underset{\sim}{\hat{\theta}})'D(\underset{\sim}{\theta}-\underset{\sim}{\hat{\theta}}), \qquad \ldots(3.33).$$

where D is a symmetric k x k non-negative definite matrix, (that is, $|D| \geq 0$). If D is positive definite, (that is, $|D| > 0$), then any non-zero error vector $(\underset{\sim}{\theta}-\underset{\sim}{\hat{\theta}})$ leads to a positive loss.

Suppose that the mean vector $E(\underset{\sim}{\Theta}/\underset{\sim}{x})$ and covariance matrix $Cov(\underset{\sim}{\Theta}/\underset{\sim}{x})$ of posterior distribution of $\underset{\sim}{\Theta}$ given exists. Thus according to equation (2.22), the Bayes estimator is that estimator which minimizes the posterior risk and is given by

$$E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}\left[[\underset{\sim}{\Theta} - \underset{\sim}{\hat{\theta}}]'D[\underset{\sim}{\Theta} - \underset{\sim}{\hat{\theta}}]/\underset{\sim}{x}\right]$$

$$=E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}\Big[[(\underset{\sim}{\Theta}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))-(\underset{\sim}{\hat{\theta}}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]'.D.[(\underset{\sim}{\Theta}$$

$$-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))-(\underset{\sim}{\hat{\theta}}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]/\underset{\sim}{x}\Big]$$

$$=E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}\left[[(\underset{\sim}{\Theta}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]'.D.[(\underset{\sim}{\Theta}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]/\underset{\sim}{x}\right] +$$

$$\left[[(\hat{\underset{\sim}{\theta}}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]'.D.[(\hat{\underset{\sim}{\theta}}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]/\underset{\sim}{x}\right]$$

$$= E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}\left[\text{ tr } [(\underset{\sim}{\Theta}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]'.D.[(\underset{\sim}{\Theta}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]/\underset{\sim}{x}\right] +$$

$$\left[[(E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x})-\hat{\underset{\sim}{\theta}})]'.D.[(E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x})-\hat{\underset{\sim}{\theta}})]\right]$$

$$= E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}\left[\text{ tr } D.[(\underset{\sim}{\Theta}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]'.[(\underset{\sim}{\Theta}-E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}))]/\underset{\sim}{x}\right] +$$

$$\left[[(E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x})-\hat{\underset{\sim}{\theta}})]'.D.[(E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x})-\hat{\underset{\sim}{\theta}})]\right]$$

$$= \text{ tr } D.\text{Cov}_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x}) + \left[[(E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x})-\hat{\underset{\sim}{\theta}})]'.D.[(E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x})-\hat{\underset{\sim}{\theta}})]\right] \quad ...(3.34).$$

Naturally, this will be minimum when

$$\hat{\underset{\sim}{\theta}} = E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x})$$

$$= \int_{\Omega_{\underset{\sim}{\theta}}} \underset{\sim}{\theta}.g(\underset{\sim}{\theta}/\underset{\sim}{x}).d\underset{\sim}{\theta}; \quad ...(3.35).$$

Therefore, posterior mean, $E_{\underset{\sim}{\Theta}/\underset{\sim}{x}}(\underset{\sim}{\Theta}/\underset{\sim}{x})$, is Bayes estimator for $\underset{\sim}{\theta}$.

3.7.1.3 Bayesian Estimation Under Absolute-Error Loss Function :

For one dimensional parameter estimation, absolute error loss function is given by

$$L(\theta-\hat{\theta}) = \omega\ |\theta-\hat{\theta}|,\ \omega > 0, \quad ...(3.36).$$

which assumes loss is proportional to the absolute value of the estimation error. Again by (3.22), Bayes estimator is that estimator which minimizes the posterior risk-$E_{\Theta/\underset{\sim}{x}}\left[\ \omega\ |\Theta - \theta|/\ \underset{\sim}{x}\right]$.

Chernoff and Moses (1959) show that the value of $\hat{\theta}$ that minimizes this posterior risk is median of the posterior distribution given $\underset{\sim}{x}$. That is, $\hat{\theta}$ is such that

$$P(\Theta \geq \hat{\theta}/\underset{\sim}{x}) \geq 0.5 \text{ and } P(\Theta \leq \hat{\theta}/\underset{\sim}{x}) \leq 0.5 \qquad \ldots(3.37).$$

Bayes risk becomes

$$r(g) = \omega\, E_{\underset{\sim}{X}}\left\{E_{\Theta/\underset{\sim}{x}}\left[\, |\Theta - \hat{\theta}| / \underset{\sim}{x}\right]\right\}, \qquad \ldots(3.38).$$

here $\hat{\theta}$ is median of the posterior distribution of Θ for given $\underset{\sim}{x}$.

3.8 Bayesian Interval Estimation :

The classical interval estimates are confidance intervals for θ and not probability intervals. Since, in Bayesian analysis θ is not a fixed constant but a r.v., Bayesian interval estimate for θ is not confidance interval but a probability interval. Assuming Θ is a single parameter for which interval estimate is desired, following definition provides *(1-γ).100% two sided Bayes probability interval*,[(1-γ).100% TBPI], for Θ.

Definition 3.5 :(1-γ).100% two sided Bayes probability interval,[(1-γ).100% TBPI] :

(θ_*, θ^*) is said to be (1-γ).100% TBPI for Θ if θ_* and θ^* satifying following posterior probabilities

$$\int_{-\infty}^{\theta_*} g(\theta/\underset{\sim}{x}).d\theta = \gamma/2 \qquad \ldots(3.39).$$

and,

$$\int_{\theta^*}^{\infty} g(\theta/\underset{\sim}{x}).d\theta = \gamma/2 \qquad \ldots(3.40).$$

Definition 3.6 :(1-γ).100% Lower Ono Sided Bayes Probability Interval,[(1-γ).100% LBPI] :

(θ_*, ∞) is said to be (1-γ).100% LBPI for Θ if θ_* satisfying following equation

$$\int_{-\infty}^{\theta_*} g(\theta/\underset{\sim}{x}).d\theta = \gamma \qquad \ldots(3.40).$$

∎

Definition 3.7 :(1-γ).100% Upper Ono Sided Bayes Probability Interval,[(1-γ).100% UBPI] :

(∞, θ^*) is said to be (1-γ).100% LBPI for Θ if θ^* satisfying following equation

$$\int_{\theta^*}^{\infty} g(\theta/\underset{\sim}{x}).d\theta = \gamma \qquad \ldots(3.41).$$

∎

Two sided Bayes pribability interval estimator provided by definition-3.5 need not be shortest interval, as every point included may not have higher probability density than every point excluded. Following definition, from Box and Tiao (1973), gives shortest interval estimate for Θ.

Definition 3.8 :Highest Posterior Density (HPD) *Region :*

Let $p(\theta/\underset{\sim}{x})$ be a posterior density function. The region R in the parameter space of Θ is called highest posterior density (HPD) region of content (1-γ) if

$$P(\Theta \in R/\underset{\sim}{x}) = 1 - \gamma \qquad \ldots(3.42).$$

$$P(\theta_1/\underset{\sim}{x}) \geq P(\theta_2/\underset{\sim}{x}), \text{ if } \theta_1 \in R \text{ and } \theta_2 \notin R. \qquad \ldots(3.43).$$

****