INTRODUCTION AND BASIC CONCEPTS

1.1 INTRODUCTION

The development of science and technology and the needs of the modern society are racing against each other. To meet the ever-increasing demands and satisfaction of the consumers, industries are trying to introduce more and more automation in their industrial processes. Thus, the complexities of industrial systems and improvement of such systems have acquired special importance in the recent years.

Generally, we come across questions like: What are the costs associated with a failure?, How frequently should maintenance or repair be carried out for keeping a system operational?, etc. Effectiveness of a system is understood to mean the suitability of the system for the fulfillment of the intended task and the efficiency of utilizing the means put into it. Reliability characteristics, such as probability of survival, mean time to system failure (MTSF) and frequency of failures are some of the measures of the effectiveness of the systems. These measures provide necessary criteria by which alternate design policies can be formulated, compared and judged and will help the system planners to select the one which best satisfies the objectives under certain techno-economic constraints. Thus, in short, the suitability of performing definite tasks is primarily determined by the quality and reliability of the system. The concepts of reliability and quality are as old as man himself.

1.2 HISTORY OF RELIABILITY AND QUALITY

Reliability, as a human attribute, has been praised for a very long time. The growth and development of the reliability theory have strong links with quality control and its development. Shewartz [1931], Dodge and Roming[1959]
laid down the theoretical basis for utilizing statistical methods in quality control of industrial products. The rate of adoption of these concepts was observed for the first time during World War I when reliability was measured as the number of accidents per hour of the flight time.

During World War II, a group in Germany was working under Wernher Von Braun developing the V-1 missiles. After the War, it was reported that the first ten V-1 missiles were fiascos. Despite attempts to provide high-quality parts and careful attention to details, all of the first missiles either exploded on the launching pad, or landed 'too soon' (in English channel). Robert Lausser, a mathematician, was called in as a consultant. His task was to analyze the missile system and he quickly derived the product probability law of series components. This theorem concerns systems functioning only if all the components are functioning and is valid under special assumptions. It says that the reliability of such a system is equal to the products of the reliability of individual components that make up to the system. If system comprises a large number of components connected in series, the system reliability may be rather low, even though the individual components have high reliability. The development continued thereafter throughout the world as increasingly more complicated products were produced in the field of engineering and technology.

With automation, the need for complicated control and safety of system becomes steadily more pressing for researchers. Practicing engineers, mathematician and statistician took interest in the study of life testing and reliability problems. United States department of defense in 1950 set the first major committee on reliability. At the same time, other countries such as Britain and Japan, began to take keen interest in the application of reliability principles of their products.
1.3 SOME BASIC RELIABILITY CHARACTERISTICS

The concept of reliability characteristics defined in many ways by different authors in the present study, the following definitions of various reliability characteristics have been used.

1.3.1 RELIABILITY

Reliability of a system is the probability that, when operating under stated environmental conditions, the system will perform its intended function adequately for a specified interval of time.

Mathematically, if \( T \) is the time till the failure of the unit occurs, then the probability that it will not fail in a given environment before time \( t \) is

\[
R(t) = P(T > t) = 1 - P(T \leq t)
\]

\[
\Rightarrow R(t) = 1 - F(t)
\]

Where, \( F(t) \) is the cumulative density function (c.d.f.) of \( T \), unreliability of the system

And

\[
R(t) + F(t) = 1
\]

Noticeable features of reliability function are:

- Function of time
- Function of environmental conditions, which may or may not vary with time.
- \( R(t) \), in general, is a decreasing function of time
- \( \lim_{t \to 0} R(t) \to 1 \) and \( \lim_{t \to \infty} R(t) \to 0 \)
- \( 0 < R(t) < 1 \)
1.3.2 HAZARD RATE FUNCTION

The measure of equipment's reliability is the infrequency with which failures occur in time. Many physical causes, individually or collectively, may be responsible for the failure of a system at any particular instant. It is not possible to isolate these physical causes and mathematically account for all of them, and therefore, the choice of failure distribution model representing the life of material or structure is still an art.

As the time passes on, the units get out worn and begin to deteriorate. These are several causes of failures of components such as:

- Careless planning, substandard equipment and raw material used, lack of proper quality control etc.
- Random or chance causes. Random failures occur quite unpredictable at random intervals and cannot be eliminated by taking necessary steps at the planning, production or inspection stage.
- Wear-out or fatigue caused by the aging of the item.

Since the item is likely to fail at any time, it is quite customary to assume that the life of the item is a random variable with a distribution function \( F(t) \) which is the probability that the item fails before time \( t \).

Many of the questions raised above can be answered if we know \( F(t) \). For example – The average life could be defined as the mean of the distribution \( F(t) \) while the probability of failure free operation between \((t_0, t_0 + t)\), given that the item was 'alive' or working at time \( t_0 \) is given by

\[
\frac{F(t_0 + t) - F(t_0)}{1 - F(t_0)} \quad \text{... (1.3.2.1)}
\]

Another very important function associated with the failure distribution \( F(t) \) is the hazard rate denoted by \( h(t) \).
Consider the probability of failure free operation with in the interval \((t, t+h)\), where \(h\) is infinitesimal, if \(f(t)\) denotes the probability density function (p.d.f.) corresponding to \(F(t)\). Then the hazard rate or the instantaneous failure rate is given by

\[
\lambda(t) = \lim_{h \to 0} \frac{F(t+h) - F(t)}{h [1 - F(t)]} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{R(t)}
\] ... (1.3.2.2)

In probability theory terms, \(r(t)\) is the conditional probability function of failure rate at instant \(t\) with the hypothesis that the unit has been functioning without failure up to that instant. Thus hazard function provides the instantaneous rate of failure at time \(t\), given that unit has survived up to the time \(t\). Equation (1.3.2.2) can be easily solved for the reliability function \(R(t)\). Therefore,

\[
R(t) = \exp \left[ - \int_0^t \lambda(u) du \right]
\]

It follows that the probability of failure free operation during the interval \((t_1, t_2)\) is expressed by

\[
R(t_1, t_2) = \exp \left[ - \int_{t_1}^{t_2} \lambda(u) du \right]
\]

There may be some probability distribution or survival function which have specific nature of hazard function with well-defined characteristics. There may be hazard functions and p.d.f.s with qualitatively different behavior. Probability model may have
- A monotonic increasing hazard function,
- A monotonic decreasing hazard function,
- A constant hazard function,
- A bathtub shaped or U-shaped hazard function

Generally three types of failures have been recognized as having a time characteristic.

(i) **INITIAL FAILURE**

Manifests itself shortly after time \( x = 0 \) and gradually begins to decrease during the initial period of operation.

Example – A good example of this type is seen in the standard human mortality table, in which it is assumed that up to the age of ten years a child can die hereditary defects, but having lived past this age, is almost certainly free of such defects.

(ii) **CHANCE FAILURE**

The chance failure, occurs the period in which a device exhibits a constant failure rate, generally lower than that prevailing during the initial period. The cause of this failure is distributed to unusually severe and unpredictable environmental conditions during the operating time of the device.

Example – In human mortality tables, it is assumed that death between the age of ten and thirty years are generally due to accidents.

(iii) **WEAR-OUT FAILURE**

The wear-out failure is associated with the gradual depletion of a material or with a accumulation of shocks, fatigue and so on.

Example – In human mortality tables, after the age of thirty years an increasing proportion of death are attributed to old age.
1.3.3 MEAN TIME TO SYSTEM FAILURE (MTSF)

The expected life, or the expected time during which a system will perform successfully, is defined as

\[ E(t) = \int_0^\infty f(t) \, dt \quad \text{... (1.3.3.1)} \]

Another convenient method of determining the expected life is given by

\[ E(t) = \int_0^\infty R(t) \, dt \quad \text{... (1.3.3.2)} \]

1.3.4 INTER-RELATIONSHIP AMONG THE RELIABILITY CHARACTERISTICS

(i) \[ f(t) = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t) \quad \text{... (1.3.4.1)} \]

(ii) \[ h(t) = \frac{f(t)}{R(t)} = \frac{1}{R(t)} \left\{ - \frac{d}{dt} R(t) \right\} = - \frac{d}{dt} \log R(t) \quad \text{... (1.3.4.2)} \]

(iii) \[ R(t) = \exp \left\{ \int_0^t h(u) \, du \right\} \quad \text{... (1.3.4.3)} \]
(vi) \[ m(t) = \frac{\int R(x)dx}{R(t)} \] \hspace{2cm} \ldots (1.3.4.4)

(v) \[ 1 + m'(t) = m(t)h(t) \] \hspace{2cm} \ldots (1.3.4.5)

(vi) \[ f(t) = h(t)\exp\left[-\int_0^t h(u)du\right] \] \hspace{2cm} \ldots (1.3.4.6)

1.4 SYSTEM

A system is defined as an arbitrary device performing an activity.

Television, Nuclear reactor, Computer, etc. are some man made systems whereas human body, solar system, etc. are natural systems. Life testing experiments can be performed either on man made systems or on natural systems to gather information. In reliability Analysis, we deal with the man-made systems while in Survival Analysis, we infer about the performance of the natural systems. Still, reliability and survival analysis can be used interchangeably.

1.4.1 PARALLEL SYSTEM

In this case, the system fails only if all its components fail. For example- if an office has n copy machines, it is possible to copy a document if at least one machine is in good working conditions.
The probability of failure of a parallel system of $n$ components is obtained as

$$P \{ \text{system failure} \} = P \left[ (y_1 = 1) \cap (y_2 = 1) \cap \ldots \cap (y_n = 1) \right]$$

Where,

$$y_i = \begin{cases} 0, & \text{if the } i^{th} \text{ component is in operative mode} \\ 1, & \text{if the } i^{th} \text{ component is in failure mode} \end{cases}$$

$$= \prod_{i=1}^{n} P_i ; \text{ if the components fail independently}$$

$$= P^n ; \text{ if in addition } P_i = P \ \forall \ i$$

### 1.4.2 LOAD-SHARING PARALLEL SYSTEMS

Many physical and non-physical systems work "in parallel" in a sense different from the case considered above. Rather than just having multiple elements connected in parallel and requiring that at least one works, the elements share the total applied "load" or demand. Parallel system of this type may function in different ways, depending on the characteristics of the components. The following are examples of two different types.

(i) The power company that interchangeably uses several power generating plants to meet total demand. In this case, the system fails if the demand exceeds the combined capacity of the power plants, i.e., the system capacity is the sum of the capacities of its components.

(ii) The rope made of several bundles that share the applied load. If the bundles are "ductile" (e.g., they are made of mild steel), they are able to redistribute the load among themselves. In this case the strength of the rope is the sum of the bundle strengths and the behavior of the system is of the same type as that of the power company in the previous example. However, if the bundles (e.g., they are made of glass), then each bundle will break as soon as capacity is reached. The load must then be carried by the surviving bundles. In this case, the strength of the rope is less than the sum of the strength of the individual components. For a system of the first type, the probability of failure is
\[ P[\text{system failure}] = P\left[ D < (C_1 + C_2 + \ldots + C_n) \right] \quad \ldots (1.4.2.1) \]

Where \( D \) is the demand and \( C_1, C_2, \ldots, C_n \) is the capacities of the Components. We shall see later in the course how to evaluate probabilities of the type in (1.4.2.1), when \( D \) and \( C_i \) are random variable.

System of the second type is more complicated to analyze, because the capacity of the system is a complicated nonlinear function of the components capacities. A way to evaluate reliability in this and other complicated cases is to use Monte Carlo simulation.

1.5 VARIOUS DISCRETE AND CONTINUOUS DISTRIBUTION USED AS LIFETIME MODELS

Numerous parametric models are used in the analysis of lifetime data and in problem related to the modeling of aging or failure process. Among univariate models, a few particular distributions, occupy a central role because of their demonstrated usefulness in a wide range of situations. Johnson and Kotz (1970) extensively catalogues mathematical and statistical properties and areas of application for various popular discrete and continuous distribution. Some frequently used lifetime models are as follows.

1.5.1 GEOMETRIC DISTRIBUTION

Consider an experiment consisting of independent trials, called Bernoulli trials, such that there are only two outcomes, \( E_1 \), and \( E_2 \). \( E_1 \) can be identified with the occurrence of a particular event and \( E_2 \) with its non-occurrence. Thus, the sample description space of this experiment is \( S = \{E_1, E_2\} \). Define a r.v. \( X_i \), such that

\[
X_i = \begin{cases} 
0 & \text{if } E_1 \text{ occurs on the } i^{th} \text{ trial} \\
1 & \text{if } E_2 \text{ occurs on the } i^{th} \text{ trial}
\end{cases}
\]
Also, let
\[ P(X_i = 0) = p \forall i \]
and
\[ P(X_i = 1) = 1 - p \forall i \]
Define another r.v. 'X' to denote the number of independent trials to the first non-occurrence. The sample description space of \( X \) is \( S = \{X : X = 1, 2, \ldots\} \) and
\[ P(X = x) = (1-p)p^{x-1}; x = 1, 2, 3, \ldots \quad (1.5.1.1) \]
The geometric distribution possesses the property similar to non-aging / Markovian / memory less property of the exponential distribution. It is the only discrete distribution that has this property.

**APPLICATIONS**

Geometric distribution can arise in the following situations:

1. In Reliability theory, suppose that every period of operation of a device is identified as a trial. If \( E_1 \) (\( E_2 \)) is identified with failure-free operation (operation with a failure). \((x-1)\) is the number of successive periods of failure-free operation, whereas \( x \) is the period at which failure occurs for the first time.

2. In the operation of time sharing computer system with fixed time slice. At the end of a time slice, the program would have completed execution with probability \( p \); thus there is a probability \( q = 1 - p(>0) \) that needs to perform more computation. The p.m.f. of r.v. denoting the number of time slices needed to complete the execution of a program is given by geometric distribution.

**1.5.2 EXPONENTIAL DISTRIBUTION**

The r.v. 'X' has an exponential distribution if it has a p.d.f. of the form:
\[ f_X(x) = \frac{\exp\left[-(x-\mu)/\theta\right]}{\theta}; (x > \mu ; \theta > 0) \quad (1.5.2.1) \]
The mathematics associated with the exponential distribution is often of a simple nature. It is often possible to obtain explicit formulae in terms of elementary functions, without troublesome quarters. For this reason, models constructed from exponential variables are sometimes used as an approximate representation of other models.

There are many situations in which exponential distribution give useful description of observed situations. One of the most widely quoted is that of events recurring “at random in time”. In applying the Monte-Carlo method, it is often required to transform random variables from standard rectangular distribution to exponential random variables.

APPLICATIONS

The exponential distribution is applied in a very wide variety of statistical procedures. Currently, among the most prominent applications are those in the field of life-testing experiments. Another application is producing usable approximate solutions to difficult distributional problems. The approximation to the distribution of a quadratic form is a case in point. This model is particularly appropriate for a situation where the failure rate appears to be more or less constant. In Reliability theory, this distribution plays very dominant role over other life-time distributions due to its forgetfulness property and simplicity in usage. It was also used in medical research by Feigel and Zelen [1965].

1.5.3 GAMMA DISTRIBUTION

Gupta and Groll [1981] proved that the gamma distribution as a natural extension of the exponential distribution and could be used as a model in life-testing experiments. The p.d.f. of the gamma distribution can be derived by considering either the time to the k-th successive arrival in a Poisson process with mean \( \lambda \) or the k-fold convolution of an exponential distribution. The p.d.f. of a two-parameter gamma distribution is given by
\begin{equation}
\Gamma(x;\lambda, k) = \frac{\lambda^k \exp(-\lambda x^k)}{\Gamma(k)} \quad ;(x,\lambda, k) > 0
\end{equation} \quad \ldots (1.5.3.1)

Here, \(\lambda\) and \(k\) are scale and shape parameters respectively and \(\Gamma(k)\) is the well-known gamma function. Though the gamma distribution does fit to wide variety of survival time data adequately, yet it is less used as a survival time model because its hazard and reliability functions are not expressible in a simple closed form. Its reliability and hazard functions involve the incomplete gamma function as:

\begin{equation}
R(t) = 1 - I(k, \lambda t)
\end{equation} \quad \ldots (1.5.3.2)

where

\begin{equation}
I(k, x) = \frac{1}{\Gamma(k)} \int_0^x u^k \exp(-u) du
\end{equation}

And

\begin{equation}
h(t) = \frac{f(t)}{R(t)}
\end{equation} \quad \ldots (1.5.3.3)

It can be shown that \(h(t)\) is monotonic decreasing (increasing) for \(k<1\) \((k>1)\) and constant for \(k = 1\). This shape parameter \(k\) is also defined as the intensity of IFR or DFR by Sharma and Rana [1990].

**APPLICATIONS**

Pearson [1900] discussed gamma distribution as the approximate distribution for "Chi-square statistics". It is also used as the approximate distribution of quadratic forms particularly positive definite quadratic forms in multi-normally distributed variables. Gamma distribution is used to represent distributions of range and quasi-ranges in random samples from a normal population. Gamma distribution share with lognormal distribution the ability to mimic closely the normal distribution (by choosing \(k\) large enough) while representing an essentially positive r.v.
1.5.4 WEIBULL DISTRIBUTION

When a theoretical model, incorporating certain assumptions about the independence, leads to an exponential distribution to represent the variation of an observable statistic, this distribution may be replaced by a Weibull distribution to allow for possible inaccuracy in the model.

DEFINITION

A random variable 'X' has Weibull distribution if there are values of the parameters $c(>0)$, $\alpha(>0)$, and $\xi_0$ such that $Y = \{(X - \xi_0)/\alpha\}^c$ has exponential distribution with p.d.f.

$$P_Y(y) = \exp(-y) ; (y > 0)$$

Then, p.d.f. of 'X' is

$$f_X(x) = cx^{-1}\{(x - \xi_0)/\alpha\}^{c-1}\exp\left[-\{(x - \xi_0)/\alpha\}\right] ; (x > \xi_0) \quad (1.5.4.1)$$

The distribution is named after Waloddi Weibull, a Swedish physicist, who used it in 1939 to represent the distribution of the breaking strength of materials.

APPLICATIONS

The Weibull distribution is perhaps the most extensively used lifetime distribution model in life-testing, reliability and quality control problems. Weibull [1951] and Berretoni [1964] advocated its applications in connection with lifetimes of many types of manufactured items. It has been used as a model with diverse type of items such as vacuum tubes [Kao, 1959], ball bearings [Lieblein and Zelen, 1956] and electrical insulation [Nelson, 1972a]. Weibull distribution is widely used in biomedical applications like in studies on time to the occurrence of tumors in human populations [Whittemore and Altshuler, 1976] or in laboratory animals [Pike, 1966, Peto et al., 1972]. It is sometimes used as a tolerance
distribution in the analysis of quantal response data. The explicit form of its cumulative distribution function is

\[ F_X(x) = 1 - \exp \left[ \left( \frac{x - \xi_0}{\alpha} \right)^c \right] \]  \hspace{1cm} \ldots (1.5.4.2)

1.6 MIXTURE OF DISTRIBUTIONS

In practice, we generally assumed that underlying population is a homogenous one with the failure time distribution given by \( F(x, \theta) \), where the form \( F \) is known but the parameter \( \theta \) is unknown. The past experience as well as experimental constraints may suggest that the assumption of homogeneity may not hold and the underlying population may consist of several subpopulations, say \( sp_1, sp_2, \ldots, sp_k \) mixed in population \( p_1, p_2, \ldots, p_k \). Further, the failure time distribution in each subpopulation is given by \( F_j(t) \), \( j = 1, 2, \ldots, k \) with p.d.f.s \( f_j(t) \) respectively. A more common situation would be when \( F_j(t) \), \( j = 1, 2, \ldots, k \) all have same from \( F \) but they differ in parameters, \( F_j(t) = F(t | \theta_j) \), \( j = 1, 2, \ldots, k \). A random sample of size \( n \) is drawn from such a population giving \( (x_1, x_2, \ldots, x_n) \) as failure times of \( n \) items included in the sample. Two different types of situations commonly arise. In one case it is possible to assign each unit to the appropriate subpopulation \( sp_j \), \( j = 1, 2, \ldots, k \) while in the other case such as information is not available. Therefore, in the first case the data would consist of the \( n \) failure times grouped according to the subpopulations,

\[ \left\{ (x_{11}, \ldots, x_{1n_1}) (x_{21}, \ldots, x_{2n_2}) \ldots (x_{k1}, \ldots, x_{kn_k}) \right\} \]

Where it is assumed that \( (n_1, n_2, \ldots, n_k) \) are the observed frequencies in the sample of the units belonging to subpopulations \( sp_1, sp_2, \ldots, sp_k \) respectively. Assuming that \( f(x | \theta_j) \) is p.d.f. corresponding to the subpopulation \( sp_j \),
straightforward computation shows that the failure time distribution of a unit is given by

\[
L(x, n; \theta, \beta) = \frac{n!}{n_1!n_2!...n_k!} \beta^{n_1} \beta^{n_2} ... \beta^{n_k} \prod_{j=1}^{k} \left[ \prod_{i=1}^{n_j} \left( x_i \mid \theta_j \right) \right] \]  \quad \cdots (1.6.1)

In the other case where we cannot assign a unit to a particular sub-population \( \pi_j \), a straightforward computation shows that the failure time distribution of a unit is given by

\[
g(x \mid \theta, \beta) = P[X \leq x] = \sum_{j=1}^{k} \beta_j f(x \mid \theta_j) \]  \quad \cdots (1.6.2)

with the corresponding p.d.f. given by

\[
g(x \mid 0, \beta) = \sum_{j=1}^{k} \beta_j f(x \mid \theta_j) \]  \quad \cdots (1.6.3)

In this case \( (x_1, x_2, ..., x_n) \), the observed failure times of \( n \) units are regarded as a sample of size \( n \) from a population with failure time distribution given by \( g(x \mid 0, \beta) \) with the corresponding p.d.f. \( g(x \mid 0, \beta) \) defined in (1.6.2) and (1.6.3) above. The likelihood of the sample is this case would be given by

\[
L(x \mid 0, \beta) = \prod_{i=1}^{n} g(x_i \mid 0, \beta) = \prod_{i=1}^{n} \left[ \sum_{j=1}^{k} \beta_j f(x_i \mid \theta_j) \right] \]  \quad \cdots (1.6.4)

Both of these model will be referred to as mixtures and our initial objective is to estimate the parameters \( (\theta_1, \theta_2, ..., \theta_k) \) with the mixing proportion \( (\beta_1, \beta_2, ..., \beta_k) \) is assumed to be known or unknown, as the case may be.

### 1.7 LIFE TESTING PLANS

In life testing experiments, either of the following situations may happen-

- A number of similar units operates in a controlled environment until all or a pre-assigned number of units fails.
• Wait for the results of the test up to uncertain time or prefix the waiting time of the test.

• Units under study are destructive in nature and can’t be used again.

• Experiment may be expensive both in terms of time and in terms of money.

• Exact life times for only a portion of the units under study are known and the remainders of the lifetimes are known to exceed certain values.

Therefore, there is the need to make the sample censored. There are various types of censoring and some of them are briefly discussed in this section.

1.7.1 TYPE-I CENSORING

Sometimes experiments are run over fixed time period in such a way that an individual’s lifetime will be known exactly only if it is less than some predetermined value. In such situations the data are said to be Type-I (or “time”) censoring. It should be noted that with Type-I censoring the number of exact lifetimes observed is random, in contrast to the case of Type-II censoring, where it is fixed.

Suppose that there are n individuals under study and that associated with the ith individual is a lifetime T_i and a fixed censoring time L_i. The T_i’s are assumed to be iid. with p.d.f. f(t) and survivor function S(t). The exact lifetime T_i of an individual will be observed only if T_i ≤ L_i. The data from such a set up can be conveniently represented by the ‘n’ pairs of random variables (t_i, δ_i), where

$$t_i = \min (T_i, L_i) \text{ and } \delta_i = \begin{cases} 1, & \text{if } T_i \leq L_i, \\ 0, & \text{if } T_i > L_i. \end{cases}$$

i.e. δ_i indicates whether the lifetime T_i is censored or not and L_i is equal to L_i if it is observed, and to L_i if it is not. The joint p.d.f. of t_i and δ_i is

$$f(t_i)^{\delta_i} S(L_i)^{1-\delta_i}$$
Time-censored / type-I censored sampling is essential in dealing with life testing experiments in which cost of experiments increases heavily with time.

1.10.2 TYPE-II CENSORING

Instead of prefixing the time to conduct an experiment if we preassign the number of items, such that the experiment is terminated as soon as these item fail, the censoring is said to be of type-II and the data are called as type-II of failure censored. Thus a type-II censored sample is one for which only r (fixed) smallest observation in a random sample of size n are recorded (1 ≤ r ≤ n) clearly, in this censoring r is fixed but time consumed up to r-th item failure is a random variables. Suppose that, the data consist of the r smallest life times $T_{(1)} ≤ T_{(2)} ≤ ... ≤ T_{(r)}$ out of a random sample of n life times $T_1, T_2, ..., T_n$ are iid and have a continuous distribution with p.d.f. $f(t)$ and reliability function $R(t)$, it follows from the general result on other statistics that the joint p.d.f. of $T_{(1)} ≤ T_{(2)} ≤ ... ≤ T_{(r)}$ is

$$\frac{n!}{(n-r)!} f(t_{(1)}) f(t_{(2)}) ... f(t_{(r)}) [R(t_{(r)})]^{n-r}$$  \hspace{1cm} (1.7.2.1)

For any given parameter model statistical inference can be based on (1.7.2.1), which gives the likelihood function and from which one can derive sampling properties of the procedures.

1.7.3 PROGRESSIVE TYPE-II CENSORING

A generalization of type-II censoring is progressive type-II censoring. In this case, the first $r_1$ failures in sample of n items are observed; then $n_1$ of the remaining "($n - n_1$)" unfailed items are removed from the experiment, leaving "($n - n_1 - n_2$)" items still present. When further $r_2$ items have failed, $n_2$ of the still un-failed items are removed, and so on. The experiment terminates after some pre arranged series of repetitions of this procedure.
1.7.4 RANDOM SAMPLING

Censoring times are often effectively random. For example, in a medical trial patients may enter the study in a more or less random fashion, according to their time of diagnosis. If the study is terminated at some prearranged date, then censoring times, that is the lengths of time from an individual's entry into the study until termination of the study, are random.

A very simple random censoring process that is often realistic is one in which each individual is assumed to have lifetime \( T \) and a censoring time \( L \), with \( T \) and \( L \) are independent continuous random variables with survivor function \( S(t) \) and \( G(t) \) respectively. Let \( \{T_i, L_i\}, i = 1, 2, ..., n \) be independent and, as in the case of Type-I censoring, define \( t_i = \min\{T_i, L_i\} \) and \( \delta_i = \begin{cases} 1 & \text{if } T_i < L_i \\ 0 & \text{if } T_i > L_i \end{cases} \).

The data from observations on \( n \) individuals consist of the pair \( (t_i, \delta_i) \), \( i = 1, 2, ..., n \). The p.d.f. of \( (t_i, \delta_i) \) is

\[
\left\{ \prod_{i=1}^{n} f(t_i)^{\delta_i} g(t_i)^{1-\delta_i} \right\} \left( \prod_{i=1}^{n} f(t_i)^{\delta_i} S(t_i)^{1-\delta_i} \right)
\]

The earlier results for Type-I censored sampling can in fact be considered as a special case of this if we allow the \( L_i \)'s to have different, degenerate distributions, each with mass at one fixed point.

1.8 PARAMETRIC / CLASSICAL INFERENCE

A model is said to be specified if the explicit form \( f_X(x; \theta_1, \theta_2, ..., \theta_k) \) is known except the values of some or all parameters \( \theta_1, \theta_2, ..., \theta_k \). Fisher [1922] in his path breaking paper titled “On the Mathematical foundations of theoretical statistics” listed the problem of classical inference into three segments, viz.,

(i) Specification of statistical / probability model
(ii) Estimation of the parameters and tests of hypotheses about the parameter

Distribution theory of statistic(s) involved.

Specification of the model means stating a statistical-mathematical model for the phenomenon under study. In fact, the correctness of model specification ensures the validity of statistical inferences whereas incorrect model specification may lead to vague inferences.

The problem of estimation of parameters and testing of hypothesis about parameters involves the choice of function (real or vector valued) of sample values, which is independent of the unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$. The estimation of parameters can be done in two ways, viz., point estimation and interval estimation. The theory of point estimation was propounded by Prof. R.A. Fisher in a series of papers.

Consider $X = (X_1, X_2, \ldots, X_n)$ be a random sample of size $n$ from $f(x; \theta)$, $\theta$ being the true value of the unknown parameter and $\hat{\theta}$ being its estimator. The consequences of estimating $\theta$ by $\hat{\theta}$ are measured in terms of loss or cost. Thus, loss is a random variable and depends on sample outcome. The accuracy and effectiveness of an estimator $\hat{\theta}$ is measured in terms of risk function i.e.

$$ R(\hat{\theta}, \theta) = \mathbb{E}(L(\hat{\theta}, \theta)) $$

The general philosophy is to choose the estimator having smaller risk among all possible estimators. There is usually no estimator which has minimum risk for all possible values of parameters but using certain principles and methods of estimation (like UMPUE, Least squares, M.L.E), it is possible to derive optimum estimators subject to unbiasedness and consistency.
1.9 BAYESIAN INFERENCE

Reliability estimation methods based on sampling theory have been found to be extremely useful for a wide variety of problems. However, there are many instances in which the classical method has been found to be less satisfactory.

Increasing insistence on cost effectiveness in reliability testing problems has had a decreasing effect on the case for consideration of sampling theory methods. If one were to consider only the use of sampling theory methods, one would be extremely limited, because of a cost and time constraints, to a very small number of samples. Such a limited sample size would result in either a very low level of confidence in the reliability estimates or imprecise estimates.

In this situation such as these, the methods based on sampling theory are frequently discarded in favor of more useful methods such as Bayesian approach.

There are two important practical benefits of a Bayesian analysis. One is the increased quality of the inferences, provided the prior information accurately reflects the true variations in parameter(s). The other is the reduction in testing requirements (test time or sample size) that often occurs in Bayesian reliability demonstration test programs. Both of these are the result of formally including additional information, in the form of the prior distribution, in the analysis.

Before the experimental data have been collected or observed, the statistician’s past experience and the knowledge will lead him to believe that \( \theta \) is more likely to lie in certain region \( \Theta \), then in others, we shall assume that the relative likelihood of the different region can be expressed in terms of a probability distribution on \( \Theta \). Thus distribution is called the prior distribution of \( \theta \), because it represents the likelihood that the true value of \( \theta \) lie in each of various region of \( \Theta \) prior to observing any values from \( f(x|\theta) \). Prior distribution \( g(\theta) \) of parameter \( \theta \) may be discrete or continuous type.
Now, suppose that \( x = (x_1, x_2, \ldots, x_n) \) be a random sample from a p.d.f. (or p.m.f.) \( f(x|\theta), \theta \in \Theta \). Then the joint p.d.f. will be
\[
f(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \quad \ldots (1.9.1)
\]
The joint p.d.f. of \( x_1, x_2, \ldots, x_n \) and \( \theta \) can now be obtained as
\[
h(x|\theta) = f(x|\theta)g(\theta) \quad \ldots (1.9.2)
\]
and the conditional p.d.f. of \( \theta \) given \( x \) denoted by \( \pi(\theta|x) \) is given by
\[
\pi(\theta|x) = \frac{f(x|\theta)g(\theta)}{\int_{\Theta} f(x|\theta)g(\theta) d\theta} \quad \ldots (1.9.3)
\]
The probability distribution over \( \Theta \) represented by the conditional p.d.f. in the above equation is called the posterior distribution of \( \theta \), after the values \( x_1, x_2, \ldots, x_n \) have been observed.

Once the posterior distribution has been obtained the posterior estimator of \( \phi(\theta) \), a function of \( \theta \), is defined as the expected value of \( \phi(\theta) \) with respect to the posterior p.d.f. i.e. the Bayes estimator of \( \phi(\theta) \) denoted by \( \phi^*(\theta) \), is given by
\[
\phi^*(\theta) = \mathbb{E}_{\pi} \left[ \phi(\theta) \mid x \right] = \int_{\Theta} \phi(\theta) \pi(\theta|x) d\theta \quad \ldots (1.9.4)
\]
In particular, Bayes estimator or posterior estimator of \( \theta \), say \( \theta^* \), becomes
\[
\theta^* = \mathbb{E}_{\pi} \left[ \theta \mid x \right] = \int_{\Theta} \theta \pi(\theta|x) d\theta \quad \ldots (1.9.5)
\]

1.10 LOSS FUNCTIONS

Consider estimating \( \theta \). Let \( \hat{\theta} \) denote an estimate of \( \theta \). The loss function, denoted by \( L(\hat{\theta}, \theta) \), is defined to be a real valued function satisfying:

(i) \( L(\hat{\theta}, \theta) \geq 0 \) for all possible estimates \( \hat{\theta} \) and all \( \theta \in \Theta \)

Chapter-1
(ii) \( L(\hat{\theta}, \hat{\theta}) = 0 \) for \( \hat{\theta} = \theta \)

\( L(\hat{\theta}, \hat{\theta}) \) equals the loss incurred if one estimates \( \theta \) to be \( \hat{\theta} \) when \( \theta \) is the true parameter value. It is a measure of the error and presumably would be greater for large error than for small error.

We now consider the following loss functions:

1.10.1 “0-1” LOSS FUNCTION

In the two action decision problem of which hypothesis testing is an example, it is typically the case that \( a_0 \) is correct if \( \theta \in \Theta_0 \) and \( a_1 \) is correct if \( \theta \in \Theta_1 \). This could correspond to testing \( H_0 : \theta \in \Theta_0 \). The loss is

\[
L(\theta, a_j) = \begin{cases} 
0 & \text{if } \theta \in \Theta_j \\
1 & \text{if } \theta \in \Theta_j (j \neq i) 
\end{cases}
\]

is called “0-1” loss function. In other words, this loss is ‘zero’ if a correct decision is made and ‘one’ if an incorrect decision is made. The interest in this loss arises from the fact that in a testing situation, the risk function of a decision rule \( \delta(x) \) is simply

\[
R(\theta, \delta(x)) = E_\theta[L(\theta, \delta(x))] = P_\theta(\delta(x) \text{ is the incorrect decision})
\]

This is either a probability of Type-I or Type-II error depending on whether \( \theta \in \Theta_0 \) or \( \theta \in \Theta_1 \). In practice, “0-1” loss will rarely be a good approximation to the true loss. More realistic losses are

\[
L(\theta, a_j) = \begin{cases} 
0 & \text{if } \theta \in \Theta_j \\
k_i(\theta) & \text{if } \theta \in \Theta_j (j \neq i) 
\end{cases}
\]

and

\[
L(\theta, a_j) = \begin{cases} 
0 & \text{if } \theta \in \Theta_i \\
k_i(\theta) & \text{if } \theta \in \Theta_i (j \neq i) 
\end{cases}
\]
This last type of loss, with \( k_1(0) \) being an increasing function of the "distance" of the true \( \theta \) from \( \Theta_1 \), is particularly reasonable, in that the harm suffered by incorrect decision will usually depend on the severity of mistake.

1.10.2 SQUARED ERROR LOSS FUNCTION:

The loss function \( L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2 \) is called squared error loss function. It was originally used in estimation problems when unbiased estimators of \( \theta \) were being considered. Since \( R(\theta, \hat{\theta}) = E_{\theta} L(\hat{\theta}, \theta) = E_{\theta} (\hat{\theta} - \theta)^2 \) would be the variance of the estimator. A second reason for the popularity of squared-error loss is due to its relationship to classical least squares theory. Finally, the use of squared error loss makes the calculations relatively straightforward and simple. The above justifications for squared error loss really have very little merit in comparison to its limitations.

Squared error loss functions are neither bounded nor concave. The convexity of it is particularly disturbing as large errors are penalized much severely. There are numerous situations in which squared error loss function is appropriate:

- Loss is symmetric in \((\theta - \hat{\theta})\) is suitable.
- Exact form of loss function is not crucial to the conclusion.
- Sample information is quite accurate in a problem with loss.

1.10.3 LINEX LOSS

A symmetric loss function assumes that positive and negative errors are equally serious. However, in some estimation problems such an assumption may be inappropriate. Canfield [1970] points out that the use of symmetric loss function may be inappropriate in the estimation of reliability function. This is in view of the fact that over-estimation of reliability function or average lifetime is
usually much more serious than under-estimation of reliability function or mean time to failure (MTTF). Also, an under-estimation of the failure rate results in more serious consequences than an over-estimation of the failure rate. This lead to the statistician to think about asymmetrical loss functions which have been proposed in statistical literature. Ferguson [1967], Zellner & Geisel [1968], Alchison & Dunsmore [1975] and Berger [1980] have considered the linear asymmetric loss function. Varian [1975] introduced the following convex loss function known is LINEX (Linear-exponential) loss function

\[ L(\Delta) = be^{q_1\Delta} - c\Delta - b \quad ; q_1, c \neq 0, b > 0 \] ... (1.10.3.1)

where \( \Delta = \hat{\theta} - \theta \). It is clear that \( L(0) = 0 \) and the minimum occurs when \( q_1b = c \), therefore, \( L(\Delta) \) can be written as

\[ L(\Delta) = b\left(e^{q_1\Delta} - q_1\Delta - 1\right) \quad ; q_1 \neq 0, b > 0 \] ... (1.10.3.2)

Where \( 'q_1' \) and \( 'b' \) are the parameters of the loss function may be defined as shape and scale respectively. This loss function has been considered by Zellner [1986], Rojo [1987]. Basu and Ebrahimi [1991] considered the \( L(\Delta) \) as

\[ L(\Delta) = b\left(e^{q_1\Delta} - q_1\Delta - 1\right) \quad ; q_1 \neq 0, b > 0 \text{ where } \Delta = \frac{\hat{\theta}}{\theta} - 1 \] ... (1.10.3.3)

and studied the Bayesian estimation under this LINEX loss function for exponential lifetime distribution. This loss function is suitable for the situations where over-estimation of \( '0' \) is more costly than its under-estimation. This loss function \( L(\Delta) \) has the following properties:

(i) for \( q_1 = 1 \), the function is quite asymmetric about zero with over-estimation being more costly than under-estimation.

(ii) for \( q_1 < 0 \), \( L(\Delta) \) rises exponentially when \( \Delta < 0 \) (under-estimation) and almost linearly when \( \Delta > 0 \) (over-estimation).
(iii) for small values of $|q_i|$. $L(\Lambda)$ is almost symmetric and not far from a squared error losses function. Indeed, on expanding $e^{q_i\Lambda} = 1 + q_i\Delta + \frac{q_i^2\Lambda^2}{2}$ or $L(\Lambda) = \frac{q_i^2\Lambda^2}{2}$ is a squared error loss function. Thus, for small values of $|q_i|$, optimal estimates are not far different from those obtained with a squared error loss function.

1.10.4 ENTROPY LOSS FUNCTION

In many practical situations, it appears to be more realistic to express the loss function in terms of the ratio $\hat{\theta}/\theta$. In this case, Calabria and Pulcini [1994] points out that a useful asymmetric loss function is the entropy loss

$$L(\delta) \propto \left\{ \delta^p - p\log_e(\delta) - 1 \right\} \quad \ldots (1.10.4.1)$$

where $\delta = \hat{\theta}/\theta$ and whose minimum occurs at $\hat{\theta} = 0$ when $p > 0$, a positive error $\left(\hat{\theta} > 0\right)$ causes more serious consequences than a negative error and vice-versa. For small $|p|$ values, the function is almost symmetric when both $\hat{\theta}$ and $\theta$ are measured in a logarithmic scale and approximately

$$L(\delta) \propto \frac{p}{2} \left\{ \log_e(\hat{\theta}) - \log_e(\theta) \right\}^2$$

Also, the loss function $L(\delta)$ has been used in Dey et al. [1987] and Dey and Liu [1992], in the original form having $p = 1$. Thus $L(\delta)$ in (1.10.4.1) can be re-written as

$$L(\delta) = b \left| \delta - \log_e(\delta) - 1 \right| ; b > 0 \text{ where } \delta = \hat{\theta}/\theta$$

1.11 PRIOR DISTRIBUTIONS

The prior distribution is an essential component of Bayesian Inference. There is no single answer to the question, "What should be the right prior?" Several solutions have been proposed by Jeffreys [1961], Raiffa and Schlaifer [1961], Degroot [1970], etc.
1.11.1 QUASI-PRIORS

For the situation where the experimenter has no prior information about the parameter \( \theta \). One may use the quasi density as given by

\[
g(\theta) = \frac{1}{\theta^d} \quad \text{if } 0 > d, \quad d \geq 0
\]

- \( d = 0 \) : Diffuse prior
- \( d = 1 \) : Non-informative prior

It is frequently happen that the natural non-informative prior is an improper prior, one, which has infinite mass, i.e.

\[
\int_{\theta \in \Theta} g(\theta) d\theta = \infty
\]

Various suggestions have been advanced for determining a non-informative prior. The most widely used method is that of Jeffreys [1961], which is to choose

\[
\prod(\theta) \propto \{I(\theta)\}^{\frac{1}{2}}
\]

As a non-informative prior, where \( I(\theta) \) is the expected Fisher information; under commonly satisfied assumptions

\[
I(\theta) = -E_\theta \left[ \frac{\partial^2 \log f(x \mid \theta)}{\partial \theta^2} \right]
\]

If \( \theta = (\theta_1, \theta_2, ..., \theta_n)^T \) is a vector. Jeffreys [1961] suggests the use of

\[
\prod(\theta) \propto \{\text{det} I(\theta)\}^{\frac{1}{2}}
\]

where \( I(\theta) \) is Fisher information matrix; \( (i,j)^{th} \) element of this matrix is

\[
I_{ij}(\theta) = -E_\theta \left[ \frac{\partial^2 \log f(x \mid \theta)}{\partial \theta_i \partial \theta_j} \right]
\]
1.11.2 NATURAL CONJUGATE PRIORS

Raiffa and Schlaifer [1961] have introduced the concept of ‘conjugate priors’ which contain a variety of prior distributions and are comparatively easy to handle. Degroot [1970] has outlined a simple and elegant methods of constructing a conjugate prior for a family of distribution \( f(x|\theta) \) which admits a sufficient statistic. The family of prior distributions \( g(\theta) \); \( \theta \in \Theta \) is called natural conjugate family if the corresponding posterior distribution belongs to the same family as \( g(\theta) \).

1.11.3 MAXIMUM ENTROPY PRIORS

In certain situations, partial information is available and outside of which it is desired to use a prior that is as non-informative as possible. A useful method of dealing with this problem is through the concept of entropy. Entropy has direct relationship to information theory and in a sense measures the amount of uncertainty inherent in the probability distributions. This method was suggested Jaynes [1968, 1983] and Rosenkrantz [1977].

Case I: \( \Theta \) is discrete

Let \( \Pi \) be a probability density on \( \Theta \)

The entropy of \( \Pi \) is defined as

\[
En(\Pi) = - \sum_{\theta} \Pi(\theta) \log(\Pi(\theta))
\]  

... (1.11.3.1)

If partial information concerning \( \theta \) is available in the form of restrictions on the \( \Pi(\theta_i) \). Then assume that

\[
\Pi(\theta_i) = \sum_k \Pi_k(\theta_i) = \mu_k \quad ; k = 1, 2, \ldots, m ; i = 1, 2, \ldots, m
\]  

... (1.11.3.2)

It seems reasonable to seek the prior, which maximizes entropy among all those priors, which satisfy the given set of restrictions. Intuitively, this results into a
prior that incorporates the available prior information but otherwise is as non-
informative as possible. The maximization of $H_n(\Pi)$ subject to (1.11.3.2) is

$$\Pi(0) = \frac{\exp \left\{ \sum_{k=1}^{m} \lambda_k g_k (0) \right\}}{\sum_{i=1}^{m} \exp \left\{ \sum_{k=1}^{m} \lambda_k g_k (0) \right\}}$$

where, $\lambda_k$'s are constants to be determined from the constraints (1.11.3.2).

Similarly, we can have the solution for continuous distribution as

$$\Pi_n(\theta) = \frac{\Pi_n(\theta) \exp \left\{ \sum_{k=1}^{m} \lambda_k g_k (\theta) \right\}}{\int_{\theta \in \Theta} \Pi_n(\theta) \exp \left\{ \sum_{k=1}^{m} \lambda_k g_k (\theta) \right\} d\theta}$$

Where, $\Pi_n(\theta)$ is the natural "invariant" non-informative prior and $\lambda_k$'s are
constants to be determined from the following constraints

$$E^H \left\{ g_k (\theta) \right\} = \int_{\theta \in \Theta} g_k (\theta) \Pi_n(\theta) d\theta = \mu_k : k = 1, 2, ..., m$$

There are several difficulties in using the maximum entropy approach to
determine prior. Firstly, usage of non-informative prior in continuous case.
Secondly, non-existence of $\Pi_n$ whenever $\Theta$ is unbounded and the specified
restrictions are specifications of fractiles of priors. This is a serious concern since
subjective knowledge about $Theta$ usually lead to fractile specifications but not to
moment specifications. Furthermore, if one attempts to subjectively specify
moments, the resulting maximum entropy prior is likely to be non-robust.

There are many physical situations, especially in spectral analysis,
available information is in the form of moment specifications and thus maximum
entropy is enormously successful.
1.12 BAYESIAN CALCULATION

Conceptually, the Bayesian paradigm is easy to implement. For any prior, $\Pi$, one need to calculate only the desired posterior feature of interest

$$E_{\pi(0|S)}[g(0)] = \frac{\int_{\Theta} g(0) f(S|\theta) \Pi(\theta) d\theta}{\int_{\Theta} f(S|\theta) \Pi(\theta) d\theta} \quad \ldots \ (1.12.1)$$

Evaluating the above integral can be quite difficult, however, especially when $\Theta$ is high dimensional. For this reason, Bayesian theory is frequently concerned with choosing $\Pi$ so as to reduce the difficulty of the calculation, while retaining essential or desirable prior feature.

In this section, discussion on Bayesian calculating methods is provided.

1.12.1 NUMERICAL INTEGRATION

Numerous integration methods like Gauss-Hermite quadrature, Simpson's rules, Weddle's rule, etc are available for calculation.

Two very important points should be kept in mind while using these methods and they are

- Don't assume that any given numerical integration method will work on a particular problem. Naylor and Smith [1982, 1983] suggested that some approximating analytical work may be needed to ensure sufficient accuracy or efficiency of the numerical integration.

- The difficulties increase rapidly with the dimension of $\Theta$. Indeed, numerical integration is rarely optimal in three or more dimensions and in such cases, Monte-Carlo integration methods become preferable.
1.12.2 MONTE-CARLO INTEGRATION

Suppose it is possible to generate an i.i.d. sequence of r.v. (θ₁, θ₂, ...), having common density h(θ).

\[ E_{h(θ)} \left( \frac{g(θ)f(x|θ)I(θ)}{h(θ)} \right) = \int_{θ∈Θ} g(θ)f(x|θ)I(θ) dθ \quad ... (1.12.2.1) \]

The key issue in Monte-Carlo integration is that of finding a suitable h(θ). On the one hand, it is desirable to choose h(θ) so that generation of random numbers (θᵢ) is inexpensive. The accuracy of the Monte-Carlo approximation can sometimes be estimated theoretically through the determination of variances of

\[ \left( \frac{f(x|θ_i)}{h(θ_i)} \right) \]. More often, however, it must be found empirically through repetition of the calculation with independent sets of r.v.'s and calculation of sample variance for equation (1.12.2.1).


1.12.3 ANALYTIC APPROXIMATIONS

The expectation in (1.12.1) can be approximated by a number of analytic methods involving Taylor's series. Though such approximations are basically "large n" approximations, yet they are often accurate for "smaller n". This will especially be true if higher order terms of the Taylor series are taken into account (Lindley [1980], and Tierney and Kadane [1985]).
1.13 A BREIF REVIEW OF LITERATURE

The research material on reliability theory may be broadly classified into two categories:

In the first category, we come across research studies, in which probabilistic models representing certain systems in practical life are visualized. These systems are then analyzed in respect of their various reliability characteristics viz. reliability, availability, mean time to system failure etc. These characteristics may be obtained under different sets of assumptions regarding operating conditions and probability laws for failure and repair of units. Studies like Dhillon and Charan Singh [1980], Govil [1983], Balagurusamy [1984] deals with such type of problems. They observed different reliability characteristics and not expected profit during a finite interval of time using the well known technique such as regeneration point technique semi-markov process and supplementary variables techniques. Generally, these operating characteristics are obtained in terms of Laplace transforms. Numerical inversion techniques can be used to evaluate these characteristics.

In the second category, we deal with the analysis of a non-negative random variable representing lifetime of a system (device). Here, lifetime of the system is assumed to have a probability distribution. Observation are recorded on the life phenomena of the system by conducting life testing experiments. Inferences are then drawn for the parameters involved in these lifetime distribution on the basis of such failure information. Since reliability characteristics of the system can be defined in terms of the parameters involved in the lifetime distribution, the drawn inferences may be transformed to these characteristics also. Studies related to such problems have been dealt with in Kapur and Lamberson [1980], Lawless [1982], Martz and Waller [1982]. The present thesis is an endeavor to extent work in reliability theory concerned with
the literature in second category where we face problems of extremely varied nature discussed as under.

Life testing experiments are costly and time-consuming phenomenon and, therefore, it should be recognized that the parameters characterizing reliability characteristics in a lifetime distributions are bound to follow some random variations due to environmental changes. Therefore, it is a factor, which should be considered with the experimental data for analyzing the reliability characteristics of the systems. Thomas Bayes [1763] introduced Bayesian inference in his famous research paper titled "An essay towards solving a problem in the Doctrine of chance". The more comprehensive studies in this regard are by Lindley [1965], Box and Tiao [1973], Savage [1962], Martz and Wallner [1982]. Bhattacharya [1967] presented the Bayesian analysis of the system reliability using many prior distributions. Some priors with their inherent statistical properties are also given in the study by Raiffa and Schlaifer [1961]. Studies like Sharma, et.al [1993, 1994, 1995] are also efforts in the same direction. But in posterior analysis, the operational experience with the complete system may not be available or very expensive to realize at the early stages of designing the system. In this regard, Kaplan, et. Al. 1989] studied about the prediction of the reliability of complete system assuming that the operational experience with the complete system is limited, non-existent or very expensive by using the information available on 'boxes'. Here, 'box' is defined as an identifiable sub-unit of a complete system. The study analyzed the behavior of various probability curves, which in turn may be used to express our degree of confidence about the complete system reliability. Recent contributions in this direction are by (Sharma and Bhutani [1992a], Sharma & Krishna [1994], and Sharma & Rana [1993]).

In the real world, there are many systems, which comprise several components arranged in series, parallel and in more complicated configurations.
The functioning of such system depends upon the performances of their components. Assuming independent and identically distributed components, the reliability analysis of these systems is well developed in the literature like Kapur and Lamberson [1980], Lewis [1987], Crowder et al. [1981]. However, in practice, there exist systems whose components are not statistically identical. Some or all of the components may have different failure-time distributions depending upon their physical properties. More so, in several situations, we need a system whose components have load-sharing attribute i.e. when any one of the components fails, its load is transmitted to the remaining surviving components.

Such types of systems are called load-sharing system models. In these models, the failure of a component within the system, increase the load shared by the surviving components. Initially, Denuit [1985] considered the load-share model for studying the reliability of a composite materials and afterwards Coleman [1957], Birnbaum and Saundes [1958] and Rosen [1964] adopted this model for analyzing the strength behavior of fiber bundles. Recently Kvam and Pena [2003], and Kim and Kvam [2004] derived methods for statistical inferences on load-share parameters in non-parametric and parametric setups respectively.

Most of the studies like Harlow and Phoenix [1978], Ross [1984], Liu [1998], Durham and Lynch [2000] dealt with the reliability estimation based on known load-share rule except for Kvam & Pena [2003], and Kim & Kvam [2004]. They derived inferences on load share parameters under equal load-share rule where the failure rates of the working components change uniformly after each failure within the system, but the magnitude of the change is unknown. Both the studies have considered the load-share system whose components having constant failure-rates.

Generally, the lifetime distributions of the systems components are assumed to be continuous, however, there exist systems whose components lifetime are measured in terms of the number of completed stock cycles. Even for
a continuous operation, involving continuous measurement of lifetime, observation made at periodic time point given rise to discrete situation, and therefore, a discrete model may be more appropriate. Yakub and Khan [1981], Patel and Gijan [1990], Patel [2003], Dilip [2004], etc considered the geometric distribution in the analysis. Further, classification of discrete lifetime have been attempted by Ray and Gupla [1992] in the multivariate setting. The studies like Hitha and Nain [1989] and Xekalake [1983] developed relationship among various reliability characteristics in the discrete bivariate setting and characterized the bivariate geometric distribution.

Mechanical/biological systems during the time of operation are always subjected to continuous environmental stresses and shocks and as such due to these stresses, one observe that the hazard-rate pattern of a device or a system over its complete lifetime is not uniform. It may follow a specific pattern up to some time epoch and change after a specific time period has elapsed. This problem to some extent has been attempted through mixed and composite lifetime distribution models like Basu [1964], Bain [1978], Lawless [1982] and Sinha [1986]. However, such an assumption seems unrealistic in view of the following facts.

There are many physical causes that individually or collectively cause the failure of a device at a time epoch. In practice, it becomes difficult to differentiate among these causes and put their effect mathematically. Thus, choosing a suitable lifetime distribution becomes a difficult task. To over come this difficulty, the studies in Mann, Scheffier and Singpurwalla [1974], and Sharma, Krishna and Singh [1997] have advocated the concept of hazard-rate model, which describes the immediate risk of failure of an individual (device) at time t, given that it has not failed up to that time. Following the above concept, Chang, Chen and Hsiung [1984], Matthews and Farewell [1982], Muller and Wang [1994], Nguyen, Rogers and Walker [1984], etc analyzed the deterministic change-point hazard-rate
model i.e. they assumed the system having a constant failure rate up to a certain length of time and changes thereafter.

1.14 THESIS AT A GLANCE

Reliability is now a well-recognized and developing branch of engineering. Since reliability study is considered essential for proper utilization and maintenance of engineering system and equipments. It has gained much importance practicing engineers and manufacture.

In classical statistical inference one applies the knowledge from the experiences to know the form of mathematical / statistical model $f(x,\theta)$ and then obtained random sample $\mathbf{x} = (x_1, x_2, ..., x_n)$ from the population. The random sample is used to draw inferences about the parameter $\theta$ or its functions. However in the Bayesian analysis implies the exploitation of suitable prior information in association with Bayes theorem. To a reliability engineer, a Bayesian approach would seem appealing because it provides a way for the formulation of a distributional form for an unknown parameter based on the prior information available to him. In reliability theory, there are several type of problems exist. To deal some of them, the present thesis has been divided into six chapters. The contents, developments and finding in each chapter are briefly discussed below.

Chapter-1 is an introductory chapter containing the basic concepts of reliability theory its inception and historical background. An introduction to well known statistical distribution with their applications, the nature of the problems of statistical inference. Bayesian approach to statistical inference and Bayesian calculations have been presented. This chapter provides an overview to different types of loss function, censoring, priors and methods of obtaining prior distributions. Towards the end of the chapter brief review of the literature in the related areas is also given.
There are many physical causes that individually or collectively cause the failure of a device at a time epoch. In practice, it becomes difficult to differentiate among these causes and put their effect mathematically. Thus, choosing a suitable lifetime distribution becomes a difficult task. To overcome this difficulty, the Chapter-2 deals with the characterization of the lifetime distribution using a hazard function. In this chapter, a probabilistic form of composite hazard-rate model is proposed. Generally, three types of failures (initial, chance and wear-out failures) have been recognized as having a time characteristic. The proposed hazard-rate model represents all the three failures, which are to be mixed in some random proportion. The reliability characteristics have been studied resulting lifetime distribution. A simulation study has been carried out for estimating the random proportion and for analyzing the various reliability characteristics both in classical and Bayesian setups.

In reliability analysis, we come across the situations, where it is neither possible nor desirable to obtain the complete sample information on the life testing experiments, as they are very costly and time consuming. In view of these facts, in chapter-3, the type-I and type-II censored failure information have been used to analyze the probabilistic hazard-rate model proposed in chapter-2. The system reliability characteristics have been analyzed both in classical and Bayesian setups. Theoretical developments are highlighted with simulated data arising from the three phases of the bathtub shaped hazard-rate model (BSHM).

In real world, there exist systems, whose components are having load-sharing attribute such as CPU of multiprocessor computers, generators in the power plant, cable in a suspension bridge and kidneys in a human body etc. More so, we encountered the systems whose component’s lifetimes are measured in terms of the number of completed cycles. In view of this, chapter-4 proposes the methods for estimation of the parameters of k-components load-sharing parallel system model that’s each component’s failure time distribution is
assumed to be geometric. The maximum likelihood estimates (M.L.E.s) of the parameters involved in the model have been obtained. To illustrate the uncertainty estimation, the variance-covariance matrices have been computed. $(1-\alpha)100\%$ confidence bounds for the parameters are also constructed. Further, past literatures have given the emphasis to the classical/empirical reliability analysis of dependent system models but to the best of our knowledge, there is no study that provides Bayesian treatment to the problem. Recognizing the fact that life testing experiments are time consuming, it seems realistic to consider the load-share parameters to be random variables. Therefore the parameters of the proposed load-share model are also estimated in the Bayesian setup.

In load-sharing systems models, the components are generally assumed to be identical. However, in practices, the system’s components are not statistically identical. Some or all of the components may have different lifetime distributions upon their physical properties. In a power plant different types of generators may be used to share the total electric load. Similarly, the cables used in a suspension bridge may be of different lengths and diameters. In the present chapter-5, we considered k-components load-share parallel systems whose some of the components are non-identical i.e. some of the components have constants failure-rates and other have linearly-increasing failure-rates. In other words, we assume that initially each component has constant failure-rate. After each failure within the system, the failure-rates of the surviving components change uniformly with unknown magnitude. However, due to continuously increase in stress on the surviving components after each failure, the system’s components experience linearly increasing failure-rates when the certain numbers of items fail. The M.L.E.s of the load-share parameters along with their simulated sample M.S.Es are obtained. To illustrate the uncertainty estimation, the variance-covariance matrices have been computed. $(1-\alpha)100\%$ confidence
bands for the parameters have also been constructed. Finally, the load-share
parameters are estimated in the Bayesian setup.

Mechanical/biological systems during the time of operation are always
subjected to continuous environmental stresses and shocks and as such due to
these stresses, one observe that the hazard-rate pattern of a device or a system
over its complete lifetime is not uniform. It may follow a specific pattern up to
some time epoch and change after a specific time period has elapsed. In view of
this, chapter-6, proposes two hazards rate models, which suit to the problems.
The estimation of the change hazard-rate time point is given in both classical and
Bayesian scenario. Theoretical developments have been highlighted with
simulated data, generated from the characterize lifetime distributions.

Programming in C++ is done for obtaining the numerical values of
estimates and given in appendix at the end of the thesis.