BAYESIAN AND ACCELERATED RELIABILITY ANALYSIS

DISSERTATION
Submitted in partial fulfilment of the Requirements for the Degree of Master of Philosophy in STATISTICS

BY
NESAR AHMAD

Under the Supervision of Dr. ARIF-UL-ISLAM

DEPARTMENT OF STATISTICS
ALIGARH MUSLIM UNIVERSITY
ALIGARH (INDIA)
1990
DEDICATED
TO
THE LOVING MEMORY OF
MY FATHER
PREFACE

Reliability theory is a complex science, related directly to engineering, physics, chemistry and economics. Reliability theory establishes the regularity of occurrence of defects in devices and methods of prediction. In recent years reliability has been formulated as the science of prediction, estimating or optimizing the probability of survival, the mean life, or, more generally, the life distribution of components or systems. The interest in Bayesian and Accelerated Reliability Theory Currently exhibited by engineers, mathematicians, economists, industrial manager and those concerned with the environmental and life sciences has stimulated the research work in this field.

In this dissertation entitled 'Bayesian and Accelerated Reliability Analysis', we have produced some basic and important results in Bayesian, non Bayesian and accelerated reliability estimation procedures.

The first chapter is an introductory part of this dissertation, which describes some fundamental aspects of statistics including probability theory, mathematical statistics and some basic concepts of stochastic process which are to be used subsequently.

The second chapter says about some of the elements of reliability concepts. Here we consider the Bayesian
Reliability analysis, life testing and accelerated life testing.

The third chapter analyses firstly a survey on Bayesian approach to reliability studies and secondly Bayesian inference in accelerated life testing and also some other Bayesian approaches, e.g., Kalman Filter approach etc.

One of the main chapter is the fourth one. The first part of this chapter is intended to be a brief exposition of the survey of accelerated life test. Many of parametric and non-parametric approaches are discussed here. This discussion includes the results based on censored as well as complete samples from accelerated tests. The second part of this chapter explores some basic assumptions to derive the life distribution of a product under accelerated life testing by progressive step when the life distribution under design constant stress follow an exponential distribution or a weibull distribution. It also discusses the special case when the acceleration equation satisfies the inverse power law and progressive stress is directly proportional to time. ML methods estimators are derived and their properties are also discussed.
The other major part of this work is chapter fifth. This chapter determines the optimal accelerated life test plans for the exponentially distributed life times under the assumptions of periodic inspection and Type I censoring.

May 28, 1990

(NESAR AHMAD)
ACKNOWLEDGEMENT

The present dissertation is being submitted to the Aligarh Muslim University, Aligarh, in partial fulfilment of the requirements for the Degree of Master of Philosophy in Statistics.

I am indebted to my supervisor Dr. Ariful Islam whose precious guidance helped me in carrying out this dissertation.

I am very much thankful to Prof. Serajur Rehman, Chairman, Department of Statistics, Aligarh for his assistance. His assistance helped me to explore new dimensions of my work. I am also thankful to Prof. S.U. Khan, Deptt. of Statistics, Aligarh, without whose blessings I would not be able to complete my M.Phil.

I am also thankful to Dr. Zaheeruddin, Dr. Hameed Khan, Mr. Abdul Aleem, Dr. S.G.A. Rizvi, Dr. Abdul Bari, Dr. Mohd. Yakoob, Dr. Jameel Hasan, Dr. Arshad Khan, Mr. Suhaib Husain and all the staff members of the department. I would also like to pay my thanks to my colleagues.

In last my special thanks to Mr. Hamendra Kumar Sharma for typing this dissertation. And very lastly I would also thank those who are jealous to me because their very jealousy inspirad me a lot.

MAY 28, 1990

(NESAR AHMAD)
CONTENTS

CHAPTER - I : FUNDAMENTALS OF PROBABILITY THEORY AND MATHEMATICAL STATISTICS. .... 1 - 21

1.1 Random variables, Probability and Probability distribution function. 1
1.2 Bayes' Theorem. 3
1.3 Important Statistical Distributions 4
   1.3.1 Gamma Distribution. 4
   1.3.2 Beta Distribution 6
   1.3.3 Multinominal Distribution 6
   1.3.4 Normal Distribution 7
   1.3.5 Exponential Distribution 8
1.4 Memoryless property of exponential distribution 8
1.5 Chebyshev Inequality. 9
1.6 Statistical Inference. 10
   1.6.1 Point Estimates and its properties 10
   1.6.2 Methods of obtaining point estimaters 12
   1.6.3 Testing of Hypothesis 13
   1.6.4 Neyman - Pearson Theorem 14
1.7 Order Statistics. 14
1.8 Stochastic Process 16
   1.8.1 Markov Process 16
1.9 Poisson Process 17
1.10 Some limit Theorems. 18
CHAPTER - II : ELEMENTS OF RELIABILITY

2.1 Basic concepts of Reliability Theory. 22
2.2 Reliability Functions 24
2.3 Statistical failure models 26
   2.3.1 The Failure Rate concept 26
   2.3.2 The Exponential Distribution 29
   2.3.3 The Weibull Distribution 30
   2.3.4 The Normal Distribution 31
   2.3.5 The logarithmic Normal Distribution 33
   2.3.6 The Gamma Distribution 34
   2.3.7 The Extreme value Distribution 35
   2.3.8 The Mixed Distribution Model 37
   2.3.9 The Composite Distribution Model 38
   2.3.10 The competing risk model 39
2.4 The Reliability of a Renewable unit 40
   2.4.1 Renewal process 40
   2.4.2 Renewal function 41
   2.4.3 Asymptotic Behaviour of a Renewal process 42
2.5 System Reliability 44
   2.5.1 Reliability of a system with independent units. 44
   2.5.2 Reliability of a system with dependent units. 46
2.6 Bayesian Reliability Analysis. 48
2.7 Life Testing. 51
2.8 Accelerated life tests. 54
CHAPTER - I

FUNDAMENTALS OF PROBABILITY THEORY AND MATHEMATICAL STATISTICS :

1.1 Random variables, Probability and probability distribution function.

1.1.1 Random Variables:

The sample space of a statistical experiment is a pair \((\Omega, \mathcal{A})\), where

i) \(\Omega\) is the set of all possible outcomes of the experiment, and

ii) \(\mathcal{A}\) is a \(\sigma\)-field of subsets of \(\Omega\).

We note that a \(\sigma\)-field is a nonempty class of subsets of \(\Omega\) that is closed under the formation of countable unions and complements and contains the null set \(\emptyset\).

A random variable is a numerical valued function defined on the sample space.

An event is any subset of the sample space. The events \(A_1\) and \(A_2\) are said to be mutually exclusive if \(A_1 \cap A_2 = \emptyset\).

1.1.2 Probability:

Let \(S\) be a sample space and let \(A\) be an event of \(S\). Consider now a number \(n\) of repeated experiments whose outcomes are described by \(S\). Let \(X\) be the number of times that \(A\) occurs in the \(n\) repeated experiments. Then the probability of event \(A\), written as \(P(A)\), is defined as

\[
P(A) \equiv \lim_{n \to \infty} \left( \frac{X}{n} \right)
\]
It is to be noted that for fixed n the quantity \( \frac{X}{N} \) is the relative frequency of the occurrence of A.

In other words, if each outcome of the sample space is equally likely. Then,

\[
P(A) = \frac{\text{number of elements of } S \text{ that corresponds to } A}{\text{total number of elements of } S}
\]

Mathematically, the probability of an event A, \( P(A) \), is a function defined on any subset A of S, such that

(i) \( P(A) \geq 0 \); 
(ii) \( P \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} P(A_i) \) for every \( i, j; i \neq j, A_i \cap A_j = \emptyset \); 
(iii) \( P(S) = 1 \)

1.1.3 Probability Distributions:

A random variable \( X \) is said to be discrete if its sample space consists of a finite or a denumerably infinite number of points. In this case, each element of \( S \) can be identified with a positive integer. Therefore,

\[ S = [x_1, x_2, \ldots, x_N] \text{ or } S = [x_1, x_2, \ldots] \]

If a function \( f_x(x_i) \) is defined on S such that

(i) \( P(X = x_i) = f_x(x_i) \); 
(ii) \( f_x(x_i) \geq 0 \); 
(iii) \( \sum_{i=1}^{N} f_x(x_i) = 1 \) or, \( \sum_{i=1}^{\infty} f_x(x_i) = 1 \)
Then \( f_x(x_1) \) is called a discrete probability distribution. If a function \( f_x(x) \) is defined on \( S \) such that

(i) \( P(a < X < b) = \int_a^b f_x(x) \, dx, \quad a < b \)

(ii) \( f_x(x) > 0; \)

(iii) \( \int_{-\infty}^{\infty} f_x(x) \, dx = 1, \)

Then \( f_x(x) \) is called a continuous probability distribution. Also, the function \( F_x(x) \) is called the distribution function and is defined as

\[
P(X \leq x) = F_x(x) = \sum_{x_1 \leq x} f_x(x_1)
\]

and for continuous random variables,

\[
F(x) = P(X \leq x) = \int_{-\infty}^{x} f(S) \, ds
\]

and also,

\[
\frac{dF(x)}{dx} = \frac{d}{dx} \int_{-\infty}^{x} f(S) \, ds = f(x)
\]

1.2 Bayes' Theorem:

Let \( \{B_1\} \) be a disjoint sequence of events such that \( P[B_i] > 0, \, i = 1, 2, \ldots, \) and \( \sum_{i=1}^{n} B_i = \varnothing, \) let \( A \)

with \( P(A) > 0. \) Then

\[
P(B_i/A) = \frac{P(A/B_i)P(B_i)}{\sum_{i=1}^{n} P(A/B_i)P(B_i)} \quad , \quad \ldots 1.2.1a
\]
1.2.1 **Bayes' Theorem for random variables**:

It is noting that conditioning the random variables $y$ be the value of $x$ does not change the relative sizes of the probabilities of those pairs $(x,y)$ that can still occur. That is to say, the probability $P(y/x)$ is proportional to $P(x,y)$. Thus

$$P(y/x) \propto P(x,y)/P(x)$$

Moreover

$$P(y) = \int P(x,y)dx = \int P(x)P(y/x)dx$$

It is clear that

$$P(y/x) = P(x,y)/P(x) = P(y)P(x/y)/P(x)$$

so that

$$p(y/x) \propto P(y)P(x/y)$$

This is of course a form of Bayes' Theorem. We note that it applies equally well if the variables $x$ and $y$ are continuous or if they are discrete. The constant of proportionality is

$$1/p(x) = 1/ \int P(y)P(x/y)dy$$

in the continuous case

$$1/P(x) = 1/ \sum_{y} P(y)P(x/y)$$

in the discrete case.

1.3 **Important Statistical Distributions**:

1.3.1 Gamma distribution

$X$ has a (one-parameter) gamma distribution with parameter $\alpha$, denoted
X \sim G(\alpha)
if it has density
\[ p(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x}, \quad 0 < x < \infty \] ...

This is simply an alternative name for the distribution we refer to as
\[ \frac{1}{2} \chi^2_{2\alpha} \]

If Y = \beta X, then Y has a two parameter gamma distribution with parameters \alpha and \beta denoted
\[ Y \sim G(\alpha, \beta) \]
and it has density
\[ p(Y) = \frac{1}{\Gamma(\alpha)\beta^\alpha} Y^{\alpha-1} e^{-Y/\beta}, \quad 0 < Y < \infty \] ...

This is simply an alternative name for the distribution we refer to as
\[ \frac{1}{2} \beta \chi^2_{2\alpha} \]

The distribution function of any variable with a gamma distribution is easily found in terms of the incomplete gamma function
\[ Y(\alpha, x) = \int_0^x z^{\alpha-1} e^{-z} \, dz \] ...

or in terms of Karl Pearson's incomplete gamma function
\[ I(u, p) = \frac{1}{\Gamma(p+1)} \int_0^u t^p e^{-t} \, dt \] ...
1.3.2 Beta distribution:

X has a beta distribution with parameter \( \alpha \) and \( \beta \), denoted 

\[ X \sim B_e(\alpha, \beta) \]

if it has density

\[ p(X) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, 0 < X < 1 \] ...

where the beta function \( B(\alpha, \beta) \) is defined by

\[ B(\alpha, \beta) = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)} \]

The mean and variance are

\[ E(X) = \frac{\alpha}{\alpha + \beta} \]  \hspace{1cm} ... \hspace{1cm} 1.3.2.2

\[ V(X) = \frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)} \]  \hspace{1cm} ... \hspace{1cm} 1.3.2.3

The distribution function of any variable with a beta distribution is easily found in terms of the incomplete beta function

\[ I_x(\alpha, \beta) = \int_0^x \frac{1}{B(\alpha, \beta)} z^{\alpha-1} (1-z)^{\beta-1} \, dz \] ...

1.3.3 Multinomial Distribution:

This distribution can be regarded as a generalisation of Binomial distribution.

When there are more than two mutually exclusive outcomes of a trial, the observations lead to multinomial distribution. Suppose \( E_1, E_2, \ldots, E_K \) are \( K \) mutually exclusive and exhaustive outcomes of a trial with the respective probabilities
The probability that $E_1$ occurs $X_1$ times, $E_2$, $x_2$ times and $E_k$, $x_k$ times in $n$ independent observations, is given by

$$p(x_1, x_2, \ldots, x_k) = C \frac{x_1^{x_1} x_2^{x_2} \cdots x_k^{x_k}}{p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}} \quad \ldots \ldots 1.3.5.1$$

where $\sum x_i = n$ and $C$ is the number of permutations of the events $E_1, E_2, \ldots, E_k$.

To determine $C$, we have to find the number of permutations of $n$ objects of which $x_1$ are of one kind, $x_2$ of another kinds, $\ldots$, and $x_k$ of the $k^{th}$ kind which is given by

$$C = \frac{n!}{x_1! x_2! \cdots x_k!}$$

Hence

$$p(x_1, \ldots, x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} \cdot p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}, \quad 0 < x_i < n$$

$$= n! \left( \prod_{i=1}^{K} \frac{x_i}{P_i} \right)^{-1} \left( \prod_{i=1}^{K} \frac{P_i}{x_i} \right) \quad \ldots \ldots 1.3.5.2$$

which is the required probability function of multinomial distribution.

1.3.4 Normal Distribution:

One of the most important distributions in the study of probability and mathematical statistics is the normal distribution.
An n X is said to have a normal distribution with parameters $\mu (\infty < \mu < \infty)$ and $\sigma (\sigma > 0)$. We will write $X \sim N(\mu, \sigma^2)$ if its pdf is given by
\[
f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-[(x-\mu)^2/2\sigma^2]} \quad -\infty < x < \infty \quad \ldots \quad \text{1.3.4.1}
\]

An r/ X is said to have a standard normal distribution if its pdf is given by
\[
\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad -\infty < x < \infty \quad \ldots \quad \text{1.3.4.2}
\]
we will write $X \sim N(0,1)$

1.4.5 Exponential Distribution:

A widely used pdf for modeling a time-to-failure variable, say $T$, is the family of exponential models defined by
\[
f(t) = \lambda e^{-\lambda t} \quad , t > 0 \quad \ldots \quad \text{1.3.5.1}
\]
\[
= 0 \quad , t < 0
\]
Note that the requirements for a valid model are not because $f(t) > 0$ for all real $t$ and $\int_{-\infty}^{\infty} f(t) dt = 1$. The function $f(t)$ assign densities to positive values of $t$.

1.4 Memoryless Property of exponential Distribution:

Suppose that $X$ has an exponential distribution given by equation (1.3.5.1). Then for any two positive integers $S$ and $t$,
\[
P(X > S + t | X > S) = P(X > t)
\]
One can say that the distribution which is satisfy the above equation has 'no memory'.

Consider,

\[
P(X > S + t / X > S) = \frac{P(X > S + t)}{P(X > S)} = e^{-\lambda(S+t) / e^{-\lambda S}} = e^{-\lambda t}
\]

Hence

\[P(X > S + t / X > S) = P(X > t)\]

Thus we have shown that the exponential distribution also has the property of having 'no memory'.

1.5 Chebyshev's Inequality:

There is a well-known inequality due to the Russian mathematician Chebyshev which will play an important role in our subsequent work. One reason that the mean and variance are said to characterize a distribution is illustrated by the Chebyshev Inequality.

Let \(X\) be a random variable (discrete or continuous with \(E(X) = \mu\) and \(V(X) = \sigma^2\). Then for any positive number \(K\) we have

\[
P(\lvert X - \mu \rvert \geq K\sigma) \leq \frac{1}{K^2}
\]

or, equivalently,

\[
P(\lvert X - \mu \rvert < K\sigma) \geq 1 - \frac{1}{K^2}
\]
1.6 Statistical Inference:

1.6.1 Point Estimates and its properties

Let $X_1, X_2, \ldots, X_n$ be a sample from $F_\theta$. Where $\theta \in \Theta \subseteq \mathbb{R}$. A statistic $T(X_1, \ldots, X_n)$ is said to be a point estimate of $\theta$ if $T$ maps $\mathbb{R}^n$ into $\Theta$.

The problem of parametric point estimation is to find an estimate $T$, for the unknown parameter $\theta$, that has some nice properties as follows:

1) **Consistency**: Let $X_1, X_2, \ldots$ be a sequence of iid rV's with common df $F_\theta$, $\theta \in \Theta$. A sequence of point estimates $T_n(X_1, \ldots, X_n) = T_n$ will be called consistent for $\theta$ if

$$T_n \xrightarrow{P} \theta \quad \text{as} \quad n \to \infty \quad \text{for each fixed} \quad \theta \in \Theta.$$

or,

Any statistic that converges stochastically to a parameter $\theta$ is called a consistent estimator of that parameter $\theta$.

Consistency is a desirable property of an estimator, and, in all cases of practical interest, maximum likelihood estimators are consistent.

2) **Invariance**: A family of probability distribution $[P_\theta : \theta \in \Theta]$ is said to be invariant under a group $G$ if for each $g \in G$ and every $\theta \in \Theta$, we can find a unique $\theta' \in \Theta$ such that the distribution of $g(X)$ is given by $P_{\theta'}$, whenever $X$ has the distribution $P_\theta$, we write $\theta' = g(\theta)$.
Remark 1. The condition that \([F_\Theta : \Theta \in \Theta]\) be invariant under \(\phi\) is the same as saying that
\[
F^*_\Theta (x_1, \ldots, x_n) = F_{g\Theta} (x_1, \ldots, x_n),
\]
where \(F^*_\Theta\) is the dt of \(g(X_1, \ldots, X_n)\) and \(F\) is the dt of \((X_1, \ldots, X_n)\).

iii) Unbiasedness: Any statistic whose mathematical expectation is equal to a parameter \(\Theta\) is called an unbiased estimator of the parameter \(\Theta\). Otherwise, the statistics is said to be biased.

iv) Sufficiency:
Let \(X = (X_1, \ldots, X_n)\) be a sample from \([F_\Theta : \Theta \in \Theta]\). A statistic \(T = T(X)\) is sufficient for \(\Theta\) or for the family of distributions \([F_\Theta : \Theta \in \Theta]\) if and only if the conditional distribution of \(X\), given \(T = t\), does not depend on \(\Theta\).

v) Completeness:
Let \([f_\Theta(x), \Theta \in \Theta]\) be a family of pdf's we say that this family is complete if
\[
E_\Theta g(X) = 0 \quad \text{for all } \Theta \in \Theta
\]
implies
\[
P_\Theta [g(X) = 0] = 1 \quad \text{for all } \Theta \in \Theta
\]
A statistic \(T(X)\) is said to be complete if the family of distributions of \(T\) is complete.
1.6.2 Methods of obtaining point estimators:

(a) Method of Moments: Let \( X_1, X_2, \ldots, X_n \) be a random sample of size \( n \) from a distribution with pdf \( f(x; \theta_1, \ldots, \theta_r) \), \( (\theta_1, \ldots, \theta_r) \in \Omega \). The expectation \( E(X_k) \) is frequently called the \( k^{th} \) moment of the distribution, \( K = 1, 2, \ldots \). The sum \( M_k = \frac{1}{n} \sum X_k^k \) is the \( k^{th} \) moment of the sample, \( K = 1, 2, 3, \ldots \). The method of moments can be described as follows: Equation \( E(X_k) \) to \( M_k \), beginning with \( K = 1 \) and continuing until there are enough equations to provide unique solutions for \( \theta_1, \theta_2, \ldots, \theta_r \), say \( h_1(M_1, M_2, \ldots) \), \( i = 1, 2, \ldots, r \), respectively.

(b) Maximum likelihood estimator:

Consider a random sample \( X_1, \ldots, X_n \) from a distribution having pdf \( f(x; \theta) \), \( \theta \in \Omega \). The joint p.d.f of \( X_1, X_2, \ldots, X_n \) is \( f(x_1; \theta)f(x_2; \theta)\ldots f(x_n; \theta) \). This joint pdf may be regarded as a function of \( \theta \). When so regarded, it is called the likelihood function \( L \) of the random sample, and we write

\[
L = (\theta, x_1, \ldots, x_n) = f(x_1; \theta)f(x_2; \theta)\ldots f(x_n; \theta), \quad \theta \in \Omega
\]

Suppose that we can find a nontrivial function of \( X_1, \ldots, X_n \), say \( U(x_1, \ldots, x_n) \), such that, when \( \theta \) is replaced by \( U(X_1, \ldots, X_n) \), the likelihood function \( L \) is a maximum. That is, \( L[U(x_1, \ldots, x_n) ; x_1, \ldots, x_n] \) is at least as great as \( L(\theta, x_1, \ldots, x_n) \) for every \( \theta \in \Omega \). Then the statistic
A statistical hypothesis is an assertion about the distribution of one or more random variables. If the statistical hypothesis completely specifies the distribution, it is called a simple statistical hypothesis, if it does not, it is called a composite statistical hypothesis.

A test of statistical hypothesis is a rule which, when the experimental sample values have been obtained, leads to a decision to accept or to reject the hypothesis under consideration.

The power function of a test of a statistical hypothesis $H_0$ against an alternative hypothesis $H_1$ is that function, defined for all distributions under consideration.

Power function = $P[X \in C/H_1]$

where $C$ is critical region.

Let $C$ denote a subset of a sample space. Then $C$ is called a best critical region of size $\alpha$ for testing the simple hypothesis $H_0 : \theta = \theta'$ against the alternative simple hypothesis $H_1 : \theta = \theta''$ if, for every subset $A$ of the sample space for which

$P [(X_1, \ldots, X_n) \in A \mid H_0] = \alpha$ ;
Let $X_1, \ldots, X_n$, where $n$ is a fixed positive integer, denote a random sample from a distribution that has pdf $f(x; \theta)$. Then the joint pdf of $X_1, \ldots, X_n$ is

$$L(\theta, x_1, \ldots, x_n) = f(x_1; \theta) f(x_2; \theta) \cdots f(x_n; \theta).$$

Let $\theta'$ and $\theta''$ be distinct fixed values of $\theta$ so that $\Omega = [\theta; \theta = \theta', \theta'']$, and let $k$ be a positive number. Let $C$ be a subset of the sample space such that

1. $\frac{L(\theta'; x_1, \ldots, x_n)}{L(\theta''; x_1, \ldots, x_n)} < k$ for each point $(x_1, x_2, \ldots, x_n) \in C$
2. $\frac{L(\theta'; x_1, \ldots, x_n)}{L(\theta''; x_1, \ldots, x_n)} \geq k$ for each point $(x_1, \ldots, x_n) \in C^*$
3. $\alpha = P \left[ x_1, \ldots, x_n \in C \mid H_0 \right].$

Then $C$ is a best critical region of size $\alpha$ for testing the simple hypothesis $H_0 : \theta = \theta'$ against the alternative simple hypothesis $H_1 : \theta = \theta''$ (1.7) order statistics:

Let $X_1, \ldots, X_n$ be a random sample from a probability density function $f_X(x, \theta)$. Suppose now that the $n$ observations are arranged in ascending order so that $X(1) \leq X(2) \leq \cdots \leq X(n)$, where $X(1)$ is the smallest
observation and \( X_n \) the largest. Then \( X_1 \) is called the first order statistic, whereas \( X_n \) is called, the \( n \)th order statistic. In general \( X_i \) is called the \( i \)th order statistic, and it has \( i-1 \) observations preceding it.

1.7.1 Distribution of the \( i \)th order statistic:

Consider the \( i \)th order statistic, \( X_i \), which has arisen from a probability density \( f_x(x, \theta) \) and a distribution function \( F_x(x, \theta) \). Also, it is assumed that \( n \) observations have been recorded, and that one needs to find the p.d.f. of \( X_i \) say \( f_{X_i}(x, \theta) \).

Let \( E \) denote the event that the \( i \)th ordered observation \( X_i \) lies between \( x \) and \( x+dx \). This implies that \( i-1 \) observations occur before \( x \) and \( n-i \) observations after \( x+dx \). Using the multinomial distribution, we have

\[
P(E) = P\left[x < X_i < x+dx\right]
= \frac{n!}{(i-1)!1!(n-i)!} [F_x(x)]^{i-1} f_x(x)dx [1-F_x(x)]^{n-i}
\]

...1.7.1.1

In the limit, \( dx \longrightarrow 0 \), we have

\[
f_{X_i}(x) = \frac{n!}{(i-1)!1!(n-i)!} [F_x(x)]^{i-1} [1-F_x(x)]^{n-i} f_x(x)
\]

...1.7.1.2
1.7.2 **Joint Distribution of the r order statistics:**

We have

\[ P[E] = P[x_1 \leq X(1) < x_1 + dx_1, \ldots, x_n \leq X(r) < x_r + dx_r] \]

where the event \( E \) here denotes the occurrence of no observations prior to \( x_1 \), the occurrence of the first observation between \( x_1 \) and \( x_1 + dx_1 \), and so on, and finally the occurrence of \( n-r \) observations after \( x_r + dx_r \). Again, using the multinomial distribution, we have

\[ P(E) = \prod_{i=0}^{n-r} f_x(x_i) dx_i \prod_{i=0}^{n-r} [1-F(x_i)]^{n-r} \]

or

\[ f(x_1, \ldots, x_r) = \prod_{i=0}^{n-r} [1-F(x_i)]^{n-r} \]

1.8 **Stochastic Process:**

A stochastic process, \([X(t); t \in T]\), is a family of random variables such that, for each \( t \) contained in the index set \( T \), \( X(t) \) is a random variable.

1.8.1 **Markov Process:**

If \([X(t), t \in T]\) is a stochastic process such that, given the value \( X(s) \), the values of \( X(t) \), \( t > S \), do not depend on the values of \( X(U) \), \( U < S \)

Then the process is said to be a Markov process.

If, for, \( t_1 < t_2 < \ldots < t_n < t \)
The process \([X(t), t \in T]\) is a Markov process.

A discrete parameter Markov process is known as a Markov chain.

### 1.9 Poisson Process:

Let \(p_n(t)\) be the probability that the random variable \(N(t)\) assumes the value \(n\) i.e.

\[ p_n(t) = P[N(t) = n] \]

This probability is a function of the time \(t\). Since the only possible values of \(n\) are \(n = 0, 1, 2, \ldots\)

\[ \sum_{n=0}^{\infty} p_n(t) = 1 \]

Thus \([p_n(t)]\) represents the probability distribution of the random variable \(N(t)\) for every value of \(t\). The family of random variables \([N(t), t \geq 0]\) is a stochastic process. The process is integral-valued. One of the most important integral-valued processes is Poisson Process.

**Postulates for Poisson Process:**

(i) Independence: \(N(t)\) is independent of the number of occurrences in an interval prior to the interval \((0, t)\), i.e. future change in \(N(t)\) are independent of the past changes.

(ii) Homogeneity in time: \(p_n(t)\) depends only on the length of the interval and is independent of where this interval is situated, i.e. \(p_n(t)\) gives the probability of the
number of occurrences in the interval \((t_1, t + t_1)\) for every \(t_1\).

(iii) Regularity: In an interval of infinitesimal length \(h\), the probability of exactly one occurrence is \(\lambda h + O(h)\) and that of more than one occurrence is of \(O(h)\). Then,

\[
\text{as } h \to 0, \frac{O(h)}{h} = 0
\]

1.10 **Some Limit Theorem:**

1.10.1 **Theorem (Poisson):** Suppose that an event \(A\) has a small probability \(p\) in each of \(n\) independent trials. Then, for large values of \(n\).

\[
P_n(m) \sim \left( \frac{\lambda^m}{m!} \right) e^{-\lambda}, \quad m = 0, 1, 2, \ldots
\]

where

\[
\lambda = np
\]

1.10.2 **The De Moivre-Laplace Integral Theorem:**

If, in a sequence of independent trials, the probability of an event \(A\) is equal to \(p\), where \(0 < p < 1\), then the probability that the number \(\mu\) of occurrences of the event \(A\) in a sequence of \(n\) trials will satisfy the inequalities.

\[
a < (\mu - np)/(npq)^{1/2} < b,
\]

where \(a\) and \(b\) are arbitrary real numbers, is for large \(n\), close to

\[
\frac{1}{(2\pi)^{1/2}} \int_{a}^{b} e^{-z^2/2} \, dz
\]
1.10.3 The Law of Large Numbers (Bernoulli's form):

Suppose that an event A has constant probability \( p \) in a sequence of independent trials. Let \( \mu \) denote the number of occurrences of the event A in \( n \) successive trials. Then, for any positive number \( \varepsilon \),

\[
P \left[ \left| \frac{\mu}{n} - p \right| > \varepsilon \right] \to 0 \text{ as } n \to \infty
\]

and Borel gives

\[
P \left[ \lim_{n \to \infty} \frac{\mu}{n} = p \right] = 1
\]

1.10.4 The Law of Large Numbers (Chebyshev's form):

Suppose that a sequence of pairwise independent random variables \( X_1, X_2, \ldots \) is such that their mean \( E(x_n) = a_n \) are finite and their variances uniformly bounded by a constant \( C [V (x_n) \leq C] \). Then, for arbitrary positive \( \varepsilon \),

\[
P \left[ \frac{1}{n} \sum_{k=1}^{n} (X_k - a_k) \leq \varepsilon \right] \geq 1 - \frac{C}{n \varepsilon^2} \text{ as } n \to \infty
\]

For this it follows that

\[
P \left[ \left| \frac{1}{n} \sum_{k=1}^{n} X_k - \frac{1}{n} \sum_{k=1}^{n} a_k \right| < \varepsilon \right] \to 1 \text{ as } n \to \infty
\]

In particular, if all the random variables have the same mean \( (a_n = a) \), then

\[
P \left[ \left| \frac{1}{n} \sum_{k=1}^{n} X_k - a \right| < \varepsilon \right] \to 1 \text{ as } n \to \infty
\]

1.10.5 The Central Limit Theorem (Lyapunov's Theorem):

We shall give a formulation of this theorem under Lindeberg's conditions.

Let \( X_1, \ldots, X_n, \ldots \) denote a given sequence of independent random variables possessing finite means and variances:
\[ a_n = E(X_n), \quad b_n^2 = V(X_n), \quad B_n^2 = \sum_{k=1}^{n} b_k^2 = V\left(\sum_{k=1}^{n} X_k\right) \]

If for some \( \tau > 0 \)

\[
\lim_{n \to \infty} \frac{1}{B_n^2} \sum_{k=1}^{n} |x - a_k| > B_n \tau \quad \Leftrightarrow \quad F_k(x) = 0
\]

...1.10.5.1

Where \( F_k(x) \) is the distribution function of \( X_k \), then, for arbitrary \( x \),

\[
\lim_{n \to \infty} P\left[ \frac{1}{B_n} \sum_{k=1}^{n} (X_k - a_k) < x \right] = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{x} e^{-z^2/2} dz
\]

...1.10.5.2

Condition (1.10.5.1) is called Lindeberg's condition.

I can be shown that, for arbitrary \( \tau > 0 \),

\[
\lim_{a \to -\infty} P\left[ \max_{1 \leq k \leq n} \left| \frac{X_k - a_k}{B_n} \right| > \tau \right] = 0
\]

Lyapunov, A.M. proceeded this theorem as

Let \( X_1, \ldots, X_n, \ldots \) denote a sequence of mutually independent random variables each with finite third moment. We introduce the notations

\[ C_k = E|X_k - a_k|^3, \quad C_n = \sum_{k=1}^{n} C_k \]

in addition to the notations introduced in the formulation of Lindeberg's theorem.

If Lyapunov's ratio

\[ L_n = C_n / B_n^2 \to 0 \]

...1.10.5.3

as \( n \to \infty \), then
Limit \( P[1/B_n \sum_{k=1}^{n} (X_k - a_k) < x] = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{x} e^{-z^2/2} \, dz \) as \( h \to \infty \).

Equation (1.10.5.3) is known as Lyapunov's condition.

1.10.6 Limiting Distributions for Maximum and Minimum Values:

Let us consider a set of \( n \) iid r.v.s \( X_1, X_2, \ldots, X_n \).

Let us define new variables \( \eta_n = \max (X_1, X_2, \ldots, X_n) \) and \( \zeta_n = \min (X_1, \ldots, X_n) \). More specifically, let the variables \( X_1, X_2, \ldots, X_n \) assume the values \( X_1^*, X_2^*, \ldots, X_n^* \). Then

\[
\eta_n = \max (X_1^*, X_2^*, \ldots, X_n^*)
\]
\[
\zeta_n = \min (X_1^*, X_2^*, \ldots, X_n^*)
\]

The distribution functions of the variables \( \eta_n \) and \( \zeta_n \) are defined by the equations:

\[
F_n(x) = P[\eta_n < x] = P[X_1 < x, X_2 < x, \ldots, X_n < x] = F^n(x)
\]

and

\[
F^*_n(x) = P[\zeta_n < x] = 1 - P[X_1 > x, X_2 > x, \ldots, X_n > x] = 1 - (1 - F(x))^n
\]

The limiting distributions for the variables \( (\eta_n - b_n) / a_n \) are exhausted by the following types:

(i) \( \Lambda(x) = e^{-e^{-x}} \)

(ii) \( \Phi_a(x) = 0 \) for \( x < 0 \) and \( e^{-x^a} \) for \( x > 0 \)

(iii) \( \Psi_a(x) = e^{-(x)^a} \) for \( x < 0 \) and 1 for \( x > 0 \).
The limiting distributions for \((\varepsilon_n - b_n) | a_n\) are exhausted by the following types:

(i) \(\Lambda^*(x) = 1 - e^{-e^x}\),

(ii) \(\Phi^*_a(x) = 1 - e^{-(x)^{-\alpha}}\) for \(x < 0\) and 1 for \(x > 0\),

(iii) \(\Psi^*_a(x) = 0\) for \(x < 0\) and \(1 - e^{-x^\alpha}\) for \(x > 0\)

The constant \(\alpha\) can assume any positive value.

*######*
CHAPTER - II

ELEMENTS OF RELIABILITY

2.1 Basic concepts of Reliability theory:

The basic concepts of reliability theory are understood by describing the relationship among them. By unit we shall mean an element, a system, a part of a system or the like. The operation of a unit means the set of all phases of its existence - transportation, maintenance, preparation for a specified use, servicing and repair. The concept of the reliability of a unit is connected in a very real way with the concept of its quality. The quality of a unit is the set of properties defining the degree of suitability of the unit for a specified use. By unit reliability we mean the ability of the unit to maintain its quality under specified conditions of use. In other words, reliability is a property which is extended in time. Reliability is determined by quality and operating conditions. We emphasize that the concept of reliability is connected with those properties that the unit had or ought to have had at the instant it was manufactured or the instant at which it was checked before use.

The problem of increasing reliability of units becomes ever more important and urgent in connection with the complex mechanization and automation of industrial processes in many fields of industry, transport, communications technology, etc. The importance of this problem is shown by the fact that insufficient reliability of units engenders great loss in their servicing, partial sloppages of equipment
and there may be accidents with considerable damage to the equipment and even human injuries. "Reliability theory is the new scientific discipline that studies the general regularity that must be maintained under design, experimentation, manufacture, acceptance and use of units in order to obtain maximal effectiveness from their use".

Failure free operation is the ability of the unit to keep its ability to function (i.e., not to have failures) throughout a specified period of time under specified condition. A failure is the partial or total loss or modification of those properties of the units in such a way that their functioning is seriously impeded or completely stopped.

The life of a unit means its capacity for extended use under the necessary technological servicing, which may include various types of repairs. At the end of the period determining the life. The unit undergoes processes associated with wear or aging the elimination of which is either impossible or economically inexpedient. For certain units, the concepts of life and failure free operation may coincide, but in general, these are independent.

For those units in which the ability to function is maintained with the aid of special renewal operations known as maintenance, and important index of reliability
is maintainability. The maintainability of a unit is its susceptibility to prediction, discovery, and elimination of failures. It is characterized by outlays in effort, time, and money made in maintenance.

Thus, the concept of reliability is shown in greater detail by the set of three concepts - failure free operation, life, and maintainability.

2.2 Reliability function:

Suppose that the unit begins to function at the instant \( t = 0 \) and that a failure occurs at the instant \( t = T \). We shall say that \( T \) is the lifetime of the unit. Let us suppose that \( T \) is a random variable with distribution given by

\[
F(t) = P \left[ T < t \right]
\]

...2.1

The function \( F(t) \) is the probability of failure of the unit prior to the instant \( t \). Let us suppose that the function \( F(t) \) is continuous and that there exists a continuous density of probability of failure.

\[
f(t) = F'(t)
\]

Thus, we have assumed that the life length of a unit \( T \) is a random variable with distribution \( F(t) \). Then let

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

\[
= P \left[ T > t \right]
\]

...2.2.2

i.e., the probability of failure free operation of the
unit during the time \( t \). The most common name for this function is 'reliability function'.

A typical form of the reliability function is shown in Fig. below. This function decrease monotonically. \( R(0) = 1 \) and \( R(t) \to 0 \) as \( t \to \infty \).

![Diagram](image)

The most important of these is the mean time of failure free operation, which is defined as the mathematical expectation of the random variable \( T \):

\[
T_o = E(T) = \int_0^\infty t \, f(t) \, dt
\]

Integrating it by parts, we get

\[
T_o = \left. t \, F(t) \right|_0^\infty - \int_0^\infty F(t) \, dt
\]

and finally we have

\[
T_o = \int_0^\infty R(t) \, dt \quad \quad \quad \quad \quad \quad \text{2.2.3}
\]

It is clear that the mean time \( T_o \) is geometrically represented by the area bounded by the coordinate axes and the curve \( R(t) \).

Another characteristic of reliability is the variance in the life length -

\[
V(T) = E(T) - T_o)^2 = E(T)^2 - [E(T)]^2
\]
2.3 STATISTICAL FAILURE MODELS:

2.3.1 The failure rate concept:

Suppose that a unit has functioned without failure up to the instant $t$. What is the probability that it will not fail during the interval $(t, t_1)$? Let us denote this probability by $R(t, t_1)$. Let $A$ denote the event that the unit functions without failure during the interval $(0, t)$ and let $B$ denote the event that it functions without failure during the interval $(t, t_1)$. Then, our probability is the conditional probability.

$$R(t, t_1) = P(B/A) = P(AB)/P(A)$$

But the event $AB$ denotes failure free operation of the unit during the interval $(0, t_1)$. Therefore,

$$R(t, t_1) = R(t_1)/R(t) \quad \ldots 2.3.1.1$$

The probability of failure during the interval $(t, t_1)$ is obviously expressed by

$$F(t, t_1) = 1 - R(t, t_1) = 1 - R(t_1)/R(t)$$

$$= [R(t) - R(t_1)]/R(t) \quad \ldots 2.3.1.2$$

Let us now set $t_1 = t + \Delta t$ and let us suppose $\Delta t$ approach zero. Then,

$$F(t, t + \Delta t) = [R(t) - R(t + \Delta t)]/R(t)$$

$$= - R'(t)/R(t) \Delta t + O(\Delta t)$$
We introduce the notation

$$\Lambda(t) = -\frac{R'(t)}{R(t)} \quad \ldots 2.3.1.3$$

For small $\Delta t$,

$$F(t, t+\Delta t) \approx \Lambda(t) \Delta t \quad \ldots 2.3.1.4$$

i.e. it determines the reliability of the unit at each given instant of time. In engineering language $\Lambda(t)$ is the probability that a unit that has functioned without failure up to the instant $t$ will fail in the interval $(t, t + \Delta t)$, or

$$\Lambda(t) = \text{Limit}_{\Delta t \to 0} \left[ \frac{\text{P(a unit of age } x \text{ will fail in the interval } (t, t + \Delta t) \text{ if it has survived up to } t)}{\Delta t} \right]$$

This function is called the Hazard rate.

Let us now consider

$$\Lambda(t) = -\frac{R'(t)}{R(t)}$$

Integrating both sides in the range of $(0, t)$, we have

$$R(t) = \text{Exp} \left[ -\int_0^t \Lambda(s) \, ds \right] \quad \ldots 2.3.1.5$$

or,

$$1-F(t) = \text{Exp} \left[ -\int_0^t \Lambda(s) \, ds \right] \quad \ldots 2.3.1.6$$

Taking derivation, we get

$$f(t) = \Lambda(t) \text{Exp} \left[ -\int_0^t \Lambda(s) \, ds \right] \quad \ldots 2.3.1.7$$

Eqn. (2.3.1.5) follows that the probability of failure free operation during the interval $(t_1, t_2)$ is expressed by

$$R(t_1, t_2) = \text{Exp} \left[ -\int_{t_1}^{t_2} \Lambda(s) \, ds \right] \quad \ldots 2.3.1.8$$
Numerous experimental data show, that for many units, the function $\lambda(t)$ has the characteristic form of the curve shown in the figure. It is clear from this graph that the entire time axis can be partitioned into three intervals. In the first of these, $\lambda(t)$ has rather high values. In any large collection of units there are always units with hidden defects and these units fail soon after they are put into operation. The first period is called the debugging or burn-in-period of defective units.

The second period is called the period of normal operation.

The last period is called the wear-out-period. In this period, the failure rate increases.
2.3.2 The Exponential Distribution:

The probability density function of the exponential distribution can be obtained either from the hazard rate concept or by considering the waiting time between arrivals in a Poisson process.

The exponential distribution is inherently associated with the Poisson process. Suppose that random 'shocks' to a device occur according to the postulates of Poisson process. Thus, the random number of shocks $X(t)$ occurring in a time interval of length $t$ is described by the Poisson distribution

$$P \left[ X(t) = n \right] = \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \quad n = 0, 1, 2, \ldots, \lambda, t > 0$$

where $\lambda$ is the rate at which the shocks occur. Further suppose that the device fails immediately upon receiving a single shock and will not fail otherwise. Let the random variable $T$ denote the failure time of the device. Thus

$$R(t) = P \left[ \text{the device survives at least to limit} \right] = P \left[ \text{no shocks occur in } (0, t) \right] = P \left[ X(t) = 0 \right] = e^{-\lambda t}$$

Thus

$$f(t) = -\frac{dR(t)}{dt} = \lambda e^{-\lambda t}, \quad t > 0, \lambda > 0$$

which is exponential distribution with parameter $\lambda$.

The same expression for the pdf of $T$ could be obtained from the hazard-rate concept. Since the assumption of a
random occurrence of shocks with parameter $\lambda$ implies a constant hazard rate, $\lambda(t) = \lambda$, for $t > 0$. Now $f(t)$ can be obtained from equation (2.3.1.7) as

$$f(t) = \lambda(t) \exp \left[ - \int_0^t \lambda(s) \, ds \right]$$

or,

$$f(t) = \lambda e^{-\lambda t}, \quad t > 0, \lambda > 0$$

and eq. (2.3.1.6) give

$$F(t) = 1 - e^{-\lambda t}, \quad t > 0, \lambda > 0$$

A more general form for the exponential distribution can be obtained if.

$$\lambda(t) = 0, \quad 0 < t < A$$

$$= \lambda, \quad t > A$$

Then,

$$f(t) = \lambda e^{-\lambda(t-A)}, \quad t > A$$

$$= 0, \quad t < A \quad \cdots 2.3.2.2$$

Often $A$ is referred to as the threshold or the shift parameter.

### 2.3.3 The Weibull Distribution:

The Weibull distribution can be derived either from the hazard rate concept or as the asymptotic distribution of the smallest order statistic from a specified probability distribution function.

Suppose we consider the hazard rate to be the power function of $t$ given by

$$\lambda(t) = \beta/\alpha \left( \frac{t-\Theta}{\alpha} \right)^{\beta-1}, \quad \alpha,\beta > 0, \Theta > 0, \ t > 0 \quad \cdots 2.3.3.1$$

$$= 0, \quad t < \Theta$$
Figure 1. The probability density function, reliability function, and hazard rate for the Weibull distribution.
Using equation (2.3.47), we have

\[ f(t) = \frac{\beta}{\alpha} \left(\frac{t-\theta}{\alpha}\right)^{\beta-1} e^{-\beta/\alpha} \int_{\theta}^{t} \left(\frac{s-\theta}{\alpha}\right)^{\beta-1} \, ds \]

Let \( \frac{s-\theta}{\alpha} = z \)

If \( s \to \theta \)

\[ Z \to 0 \]

\[ ds = \alpha \, dz \]

and if \( s \to t, \, Z \to \frac{t-\theta}{\alpha} \)

\[ f(t) = \frac{\beta}{\alpha} \left(\frac{t-\theta}{\alpha}\right)^{\beta-1} e^{-\beta/\alpha} \left(\frac{t-\theta}{\alpha}\right)^{\beta} \quad t > \theta \ldots 2.3.3.2 \]

Similarly, using (2.3.1.5), the corresponding Weibull reliability function becomes

\[ R(t) = e^{-\left(\frac{t-\theta}{\alpha}\right)^{\beta}} \quad \ldots 2.3.3.3 \]

In figure 4c, the hazard rate is decreasing (increasing) in \( t - \theta \) if \( \beta < 1 (\beta > 1) \), and is constant if \( \beta = 1 \), when \( \beta = 1 \) the Weibull distribution specializes to the exponential distribution, and when \( \beta = 2 \) the resulting distribution is known as the Rayleigh distribution.

Since \( \beta \) controls the shape of the distribution, it is called the shape parameter, while \( \alpha \) and \( \theta \) are usually referred to as the scale and location parameters, respectively. Again, \( \theta \) corresponds to a period of guaranteed life that is not present in many applications, thus \( \theta = 0 \) in such instances.

2.3.4 The Normal Distribution:

Sudden failures of a random nature are usually described by an exponential law. On the other hand, failures that
arise as the result of wear and tear, of irreversible physico-chemical changes in the physical parameters of the unit do not obey an exponential law. These failures, known as wear-out failures, are described by a normal law.

For a normal reliability law, the reliability function is of the form

$$R(t) = \frac{\frac{1}{\sqrt{2\pi}} \int_{(t-T_0)/\sigma}^{\infty} e^{-x^2/2} \, dx}{\frac{1}{\sqrt{2\pi}} \int_{-T_0/\sigma}^{\infty} e^{-x^2/2} \, dx} \quad \text{...2.3.4.1}$$

and since, $\sigma \ll T_0$, we can write this equation as follows

$$R(t) = \frac{1}{\sqrt{2\pi}} \int_{(t-T_0)/\sigma}^{\infty} e^{-x^2/2} \, dx \quad \text{...2.3.4.2}$$

where, $T_0$ is the mean life length and $\sigma^2 = V(T)$. One can show that the failure rate $\lambda(t)$ for a normal law is of the form (see fig.)

If increases monotonically and, after $T_0$, it begins, to approach as an asymptote the line $(t - T_0)/\sigma$. Also, the failure rate is not a significant characteristic for large values of $y$. 
2.3.5 The Logarithmic Normal Distribution:

The logarithmic normal distribution implies that the logarithms of the life times are normally distributed, hence it can be easily derived by a simple logarithmic transformation. The hazard rate of the logarithmic normal distribution as a function of time is an increasing function followed by a decreasing function, and can be shown to approach zero for large life times and at the initial time. For this reason the derivation of the logarithmic normal distribution from the hazard rate is difficult. Its derivation by a logarithmic transformation is given below.

Let $T$ be the time to failure random variable of a device, and let $X = \log_e T$ be distributed normally with parameters $\mu$ and $\sigma$.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-1/2 \left( \frac{x - \mu}{\sigma} \right)^2}, \quad -\infty < x < \infty$$

It follows from the above that the p.d.f. of $T$, $f(t)$ is given by

$$f(t) = \frac{1}{\sigma t \sqrt{2\pi}} e^{-1/2\sigma^2 (\log t - \mu)^2}, \quad 0 < t < \infty$$

$$\quad \quad -\infty < \mu < \infty$$

$$\quad \quad \quad \sigma^2 > 0 \quad \cdots 2.3.5.1$$

This is the logarithmic normal pdf for $T$.

The reliability function for this law is

$$R(t) = \frac{1}{\sqrt{2\pi}} \cdot \int_{\log(t/T_0)/\sigma}^{\infty} e^{-x^2/2} \, dx \quad \cdots 2.3.5.2$$
2.3.6 The Gamma Distribution:

The gamma distribution is a natural extension of the exponential distribution and is used as a failure time model. It is also used as a prior distribution in Bayesian reliability analysis.

The gamma distribution can be derived by considering the time to the occurrence of the \( n \)th event in a poisson process [Mann, Schafer, and Singpurwalla, (1974)]. For example if the time \( T_1 \) between successive failures of a system has an exponential distribution, then \( T = T_1 + \ldots + T_n \). The cumulative time to the \( n \)th failure, follows a gamma distribution with scale parameter and shape parameter \( n \). The gamma distribution is thus said to be the \( n \)-fold convolution of an exponential distribution [Feller 1965].

There is an alternative way of gamma failure time distribution. Consider a situation in which a system operates in an environment where shocks occur according to a poisson process. Further, suppose that the system fails upon receipt of exactly \( n \) shocks and not before. The system failure time \( T \), which denotes the random time of occurrence of the \( n \)th shock, follows a gamma distribution with parameter \( \lambda \) and \( n \).

The pdf of a gamma distribution with parameters \( \alpha \) and \( \beta \) is given by

\[
f(t) = \frac{1}{[\Gamma(\alpha)]} t^{\alpha-1} e^{-t/\beta} \quad ; \alpha, \beta, t > 0 \quad \ldots\ldots 2.3.6.1
\]
where \( \Gamma(\alpha) \) denotes the gamma function evaluated at \( \alpha \).

The gamma distribution will be denoted by \( G(\alpha, \beta) \). The parameters \( \alpha \) and \( \beta \) are referred to as the shape and scale parameters of the \( G(\alpha, \beta) \) distribution.

For \( \alpha = n/2 \) and \( \beta = 2 \), we obtain a \( \chi^2 \) distribution and for \( \alpha = 1 \), \( G(\alpha, \beta) \) reduces to the exponential distribution with parameter \( 1/\beta \).

The reliability function are not available in closed form unless \( \alpha \) happens to be an integer. If may be expressed in terms of the incomplete gamma function \( \Gamma(\alpha, z) \) defined by

\[
\Gamma(\alpha, z) = \int_0^z y^{\alpha-1} e^{-y} \, dy, \quad \alpha > 0
\]

In this term, the reliability function is given by

\[
R(t) = \frac{\Gamma(\alpha) - \Gamma(\alpha, t/\beta)}{\Gamma(\alpha)}
\]

which, if \( \alpha \) is an integer, becomes

\[
R(t) = \frac{t^{\alpha-1} e^{-t/\beta}}{\beta^\alpha \Gamma(\alpha)}
\]

Mean and variance of (2.3.6.1) is given as

\[
\mu = E(t) = \alpha \beta \\
\sigma^2 = \text{Var}(t) = \alpha \beta^2
\]

2.3.7 The Extreme Value Distribution:

Let us consider a random sample of size \( n \) from an infinite population having a cumulative distribution function \( F(x) \) where \( X \) is a continuous random variable
Let the sample be denoted as \( x_1, x_2, \ldots, x_n \). Define the random variable

\[ y_n = \min (X_1, \ldots, X_n) \]

The random variable \( y_n \) is termed the smallest extreme value.

Since material or equipment failure is related to the weakest or the weakest component, the extreme value distribution for the smallest value is applicable in reliability work. The smallest extreme value distribution will be considered here:

The cumulative distribution for \( y_n \) is given by

\[ P(y_n > y) = P[(X_1 > y) \cap (X_2 > y) \cap \ldots \cap (X_n > y)] \]

\[ \Rightarrow P(y_n > y) = \prod_{i=1}^{n} P(X_i > y) \]

or,

\[ P(y_n > y) = [1 - F(y)]^n \]

and then the cumulative distribution for \( y_n \) is

\[ G_n(y) = 1 - (1 - F(y))^n, \quad -\infty < y < \infty \quad \ldots 2.3.7.1 \]

Hence the pdf is

\[ g_n(y) = nf(y) [1 - F(y)]^{n-1}, \quad -\infty < y < \infty \quad \ldots 2.3.7.2 \]

This is also called the pdf for the first order statistic in a sample of size \( n \).

If \( f(x) \) is integrable, then \( G_n(y) \) is easily obtained, however, \( f(x) \) is not always integrable and this had led to the study of \( G_n(y) \) as \( n \) becomes large.
Define the random variable $U_n$ as

$$U_n = n F(y_n) \quad \ldots 2.3.7.3$$

Here $F(x)$ is the cdf for $x$ and since $0 \leq F(x) \leq 1$, then

$$0 \leq U_n \leq n$$

Now, the cdf of $U_n$ is

$$H_n(u) = P(U_n \leq u) = P(nF(y_n) \leq u) = P[y_n \leq F^{-1}(u/n)]$$

$$= G_n[F^{-1}(u/n)]$$

From equation (2.3.7.1) we have

$$H_n(u) = 1 - (1 - u/n)^n, \quad 0 \leq u \leq n \quad \ldots 2.3.7.4$$

As $n \to \infty$ this becomes

$$H(u) = 1 - e^{-u}, \quad u \geq 0 \quad \ldots 2.3.7.5$$

and

$$h(u) = H'(u) = e^{-u}, \quad u \geq 0 \quad \ldots 2.3.7.6$$

Thus the limiting distribution of the smallest extreme value is given by the distribution of $y$.

2.3.8 The Mixed Distribution Model:

It has been a common practice to assume that distribution that are mixed belong to the same family but differ in the values of their parameters. The mixed normal distribution considered by Cohen is claimed to be applicable to the study of wind velocities and physical dimensions of mass-produced items. The mixed Weibull distribution discussed by Kao is useful in reliability studies, especially those involving electron tubes.

Let $F_{T_i}(t)$ be the cumulative distribution function of a random variable $T_i, \ i = 1, 2, \ldots k$. Then a $k$-fold mixed c.d.f. is defined as
\[ F_T(t) = \sum_{i=1}^{k} p_i F_{T_i}(t) \quad 0 \leq p_i \leq 1, \text{ and } \sum_{i=1}^{k} p_i = 1 \]

Often \( F_{T_i}(t) \) is referred to as the \( i \)th subpopulation in c.d.f form, and the \( p_i \) are called the mix parameters.

Alternatively, a \( k \)-fold mixed probability density function is given by
\[ f_T(t) = \sum_{i=1}^{k} p_i f_{T_i}(t) \]

where \( f_{T_i}(t) \) is the \( i \)th subpopulation in pdf form.

### 2.3.9 The Composite Distribution Model:

The main reason for considering a composite distribution model is that it can sometimes provide flexibility in fitting and explaining failure data.

An \( r \)-component composite c.d.f. is defined as
\[ F_T(t) = F_j(t), \quad S_j \leq t \leq S_{j+1}, \quad j = 0, 1, 2, \ldots r. \]

Here \( F_j(t) \) is called the \( j \)th component of a composite distribution in c.d.f. form. The parameters \( S_j \) are termed the partition parameters. It is clear than an \( r \)-component composite weibull distribution has \( r-1 \) partition parameters. An \( r \)-component composite p.d.f. is simply
\[ f_T(t) = f_j(t), \quad S_j \leq t \leq S_{j+1}, \quad j = 0, 1, 2, \ldots r \]

Where \( f_j(t) \) is the derivative of \( F_j(t) \).
2.3.10 The Completing Risk Model:

Suppose that a device exhibits $K$ modes (risks) of failure, $m_1, m_2, \ldots, m_k$, and that a random life time on this item occurs as follows: when the device begins operation, each failure mode simultaneously generates a random life time that is independent of the other modes. Thus, in effect, $K$ lifetimes, denoted by $T_1, T_2, \ldots, T_K$, simultaneously begin, lifetime $T_i$ corresponds to the $i$th mode of failure. Failure of the device occurs as soon as any one of the lifetimes, say $T_i$, is realized. In effect, if the life length of the device is denoted by a random variable $T$, then

$$T = \min (T_1, T_2, \ldots, T_K) = T(l).$$

If $F_{T_i}(t)$ is the cumulative distribution function of $T_i$, the c.d.f. of $T$, $F_T(t)$, is given by

$$F_T(t) = 1 - \prod_{i=1}^{k} [1 - F_{T_i}(t)]$$ \hspace{1cm} \text{....2.3.10.1}

The above derivation of the completing risk model not only is independent of the functional form of the $F_{T_i}(t)$, but also allows for the $F_{T_i}(t)$ to be all different. In effect, this means that each failure mode can have any failure distribution and that all the failure distribution need not be alike.

Consider a device an which $k$ risks are jointly but independently operating. If only the $i$th risk, with a risk-specific hazard rate $\lambda_i(s)$ were effective, the
probability that the device will survive to time \( t \) is
\[
1 - F_T(t) = \exp \left[ - \int_0^t \lambda_1(s) \, ds \right] \quad \ldots 2.3.10.2
\]
where \( \int_0^t \lambda_1(s) \, ds \) is the cumulative hazard due to risk 1, at time \( t \). From eqn.(2.3.10.1) it is clear that the probability that the device will survive all the \( k \) risks is
\[
1 - F_T(t) = \prod_{i=1}^k \exp \left[ - \int_0^t \lambda_i(s) \, ds \right] 
= \exp \left[ - \int_0^t \sum_{i=1}^k \lambda_i(s) \, ds \right] \quad \ldots 2.3.10.3
\]

This equation leads to the inference that the total hazard to the device at time \( S \), say \( \Lambda(s) \), is the sum of the \( k \) independent risk-specific hazards at time \( S \), that is
\[
\Lambda(s) = \lambda_1(s) + \lambda_2(s) + \ldots + \lambda_k(s).
\]

If it is assumed that the risk-specific hazard rates are constant over the period of observation i.e.
\[
\lambda_1(s) = \lambda_1 \quad \forall \, 1 \text{ and } S > 0, \text{ then} \\
\lambda(s) = \lambda_1 + \lambda_2 + \ldots + \lambda_k = \lambda \text{ (say)}
\]
and
\[
1 - F_T(t) = \exp (-\lambda t) \quad \ldots 2.3.10.4
\]

2.4 THE RELIABILITY OF A 'RENEWABLE' UNIT:

2.4.1 **Renewal Process**:

In this process we assume that, after a failure, the unit is renewed. This renewal can assume various forms: it can be replaced with a new unit that is
identical to it or it can be subjected to maintenance that completely restores all its original properties. We shall assume that as soon as a unit fails it is renewed instantaneously. Suppose that the unit begins operating at the instant \( t = 0 \) and continues operating for a random period of time \( T_1 \) and then fails. At that instant, it is replaced with a new unit. Which operates for a length of time \( T_2 \), then fails and is replaced with a third unit. This process is continued indefinitely.

It is natural to assume that the life lengths \( T_1, T_2, \ldots \) of the units are independent. The random times \( T_1, T_2, \ldots \) have the same distribution \( F(t) \).

\[
F(t) = P(T_n < t)
\]

It is clear from the above figure that the instant of failures or renewals

\[
t_1 = T_1, \ t_2 = T_1 + T_2, \ldots, \ t_n = T_1 + T_2 + \ldots + T_n, \ldots
\]

Constitutes a random flow, which we shall call a renewal process.

2.4.2 Renewal function:

Let \( \Upsilon(t) \) = number of failures that take place in the time \( t \). Obviously \( \Upsilon(t) \) is a random variable. Let us find the distribution of \( \Upsilon(t) \). We note that

\[
P[\Upsilon(t) > n] = P[t_n < t] = P[T_1 + T_2 + \ldots + T_n < t] = F_n(t)
\]

...2.4.2.1
where the functions \( F_n(t) \) are the distribution laws of the \( t_n \) and are defined by

\[
F_n(t) = \int_0^t F_{n-1} (t - T) \, dF(T), \quad F_1(t) = F(t)
\]

Equation (2.4.2.1) implies that

\[
R_n(t) = P[ \gamma(t) = n] = F_n(t) - F_{n+1}(t) \quad \ldots \quad 2.4.2.2
\]

In particular,

\[
R_0(t) = 1 - F(t).
\]

The renewal function \( H(t) \), is defined as the mean number of failures that occur up to the instant \( t \). By using equation (2.4.2.2.), we get

\[
H(t) = E[\gamma(t)] = \sum_{n=1}^{\infty} n R_n(t)
\]

\[
= \sum_{n=1}^{\infty} n F_n(t) - \sum_{n=2}^{\infty} (n-1) F_n(t)
\]

\[
= \sum_{n=1}^{\infty} F_n(t) \quad \ldots \quad 2.4.2.3
\]

Also, \( h(t) = H'(t) \)

The function \( h(t) \) is called the renewal density. It is equal to the mean number of failures that take place in a unit interval beginning at the instant \( t \). From eqn. (2.4.2.3), we have

\[
h(t) = \sum_{n=1}^{\infty} f_n(t) \quad \text{where} \quad f_n(t) = F'_n(t) \quad \ldots \quad 2.4.2.4
\]

\[\textbf{2.4.3 Asymptotic Behaviour of a Renewal Process :}\]

Here we consider the behaviour of the process on the subintervals preceding by a large number of failures.
Therefore we study the asymptotic behaviour of the process as \( t \to \infty \).

We know that from eqn. (2.4.2.1)

\[ P \left( \forall (t) > n \right) = P \left( T_1 + T_2 + \ldots + T_n < t \right). \]

Since the variables \( T_i \) are identically distributed and have a finite variances \( \sigma^2 \), the sequence \( \left\{ \xi_n \right\} \) of fractions

\[ \xi_n = \frac{T_1 + T_2 + \ldots + T_n + n T_0}{\sigma(n)^{1/2}} \]

\( \Rightarrow \xi_n \sim N(0,1) \) as \( n \to \infty \).

Suppose now that \( t \to \infty \) and

\[ h = T_0 h^2 + z_n(t)^{1/2} \]

where \( z_n \) are chosen in such a way that \( z_n \to Z \) and \( n \) is an integer.

Let us rewrite Eqn. (2.4.2.1) in the form

\[ P \left( \frac{\forall (t) - (t/T_0)}{(t)^{1/2}} > Z_n \right) = P \left( \xi < \frac{-T_0 z_n(t)^{1/2}}{\sigma((T_0)^{1/2})^{1/2}} \right) \]

As \( t \to \infty \), we obtain

\[ \lim_{t \to \infty} P \left[ \frac{\forall (t) - (t/T_0)}{(t)^{1/2}} > Z_n \right] = P \left( \xi < \frac{-T_0^{3/2} Z}{\sigma} \right) \]

Let us make the substitution \( x = T_0^{3/2} z/\sigma \). Then,

\[ \lim_{t \to \infty} P \left[ \frac{\forall (t) - (t/T_0)}{\sigma(t)^{1/2} T_0^{3/2}} > x \right] = P \left[ \xi < -x \right] \]

\[ = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} e^{-t^2/2} dt \]

\[ \ldots 2.4.3.2 \]
From this it follows that the random variable $\sqrt{t}$ is asymptotically normal with mean 

$$E[\sqrt{t}] \sim t/T_0$$

and with variance

$$V[\sqrt{t}] \sim \sigma^2 t/T_0^3$$

This fact enables us to estimate in a simple manner and rather accurately the number of failures in a long interval of time.

2.5 SYSTEM RELIABILITY :

2.5.1 The Reliability of a system with Independent units :

By 'System' we mean an arbitrary device consisting of parts the reliabilities of which are given. We shall call these parts 'units'. We assume that the units fail independently of each other, i.e., that the failure of a unit does not change the reliability of other units.

Suppose that the system consists of $n$ units the reliability functions of which are given as $R_1(t)$, $R_2(t)$, ..., $R_n(t)$.

Our problem consists in expressing the reliability function $R(t)$ of the system in terms of the reliability functions of the units.

Case 1 : Series System :

We say that $n$ units in a system are connected in series in a reliability sense, if failure of an arbitrary element causes failure of the entire systems.
Let $R(t)$ be the reliability function of the system. Since the units are independent and for failure free operation of the system for a period of time $t$, it is necessary that each unit operate without failure during that period. Therefore,

$$R(t) = R_1(t)R_2(t)\ldots R_n(t)$$  \hspace{1cm} (2.5.1.1)

which is known as the product rule in Reliability. It is observed that the system reliability will always be use than or equal to the least reliable component. Thus,

$$R(t) \leq \min_t [R_i(t)]$$

Let us express the reliability functions in terms of failure rates.

$$e^{-\int_0^t \lambda(s)ds} = e^{-\int_0^t \lambda_1(s)ds} - \int_0^t \lambda_2(s)ds \ldots - \int_0^t \lambda_n(s)ds$$

$$\Rightarrow \lambda(t) = \lambda_1(t) + \lambda_2(t) + \ldots + \lambda_n(t)$$  \hspace{1cm} (2.5.1.2)

i.e., the failure rates are added.

Suppose that $n_i$ units have reliability function $R_i(t)$ in the $i^{th}$ set, $i = 1, 2, \ldots, S$, in a complex system. Then, equation (2.5.1.1) and (2.5.1.2) become

$$R(t) = [R_1(t)]^{n_1} [R_2(t)]^{n_2} \ldots [R_S(t)]^{n_S}$$

$$\lambda(t) = n_1 \lambda_1(t) + n_2 \lambda_2(t) + \ldots + n_S \lambda_S(t)$$

Case-II : Parallel System :

We say that the units in a system are connected in
paralle if the system fails only when all the units of the system fail. Thus
\[
F(t) = P(T \leq t) = P(T_1 \leq t) \cdot P(T_2 \leq t) \cdots P(T_n \leq t),
\]
\[
= \prod_{i=1}^{n} F_i(t) = \prod_{i=1}^{n} [1 - R_i(t)]
\]
or,
\[
R(t) = 1 - F(t) = 1 - \prod_{i=1}^{n} [1 - R_i(t)]
\]
In case of exponential distributions
\[
R(t) = 1 - \prod_{i=1}^{n} (1 - e^{-\lambda_i t})
\]
and thus, \( T \) does not have an exp. distribution, i.e., if the reliability of each unit obeys an exp. law, the reliability of the system does not obey this law.

2.5.2 The Reliability of a system with Dependent units:

Up to now, we have always assumed that the units in a system are independent in the sense of reliability. It is a stringent restriction. Failures of a certain units can influence in a very significant way the reliability of other units. In the case of connection in parallel, failures of certain units lead to increase in the functional load applied to the units still in operation and reliability of these units drops. Suppose that a system consists of \( n \) identical units connected in parallel. Let us assume that the failure rate of each unit is independent of time but
depends on the number of units that have not failed. If, at a given instant, \( K \) units are in operation, then the failure rate of each of them is equal to \( \lambda_k \). Let us denote by \( R_k(t) \) the probability that exactly \( k \) units are in operation at the instant \( t \). In this case, we say that the system is in the \( k \)th state. It can leave this state and enter the \((k - 1)\)st state in an infinitesimally small time \( \Delta t \) with probability

\[
K \lambda_k \Delta t + O(\Delta t)
\]

and it can remain in the state \( K \) with probability

\[
1 - K \lambda_k \Delta t + O(\Delta t).
\]

Comparing the state of the system at two infinitesimally close instants \( t \) and \( t + \Delta t \), we obtain from the formula for total probabilities

\[
R_k(t + \Delta t) = [ (k + 1) \lambda_{k+1} \Delta t + O(\Delta t)] R_{k+1}(t) + [1 - (\lambda_k \Delta t + O(\Delta t))] R_k(t) + O(\Delta t)
\]

Dividing both sides of this equation by \( \Delta t \) and taking the limit as \( \Delta t \to 0 \), we obtain

\[
R'_k(t) = (k + 1) \lambda_{k+1} R_{k+1}(t) - K \lambda_k R_k(t),
\]

\( K = 0, 1, 2, \ldots, n-1 \)

\[\ldots.2.5.2.1\]

For \( K = n \), we obtain analogously

\[
R'_n(t) = -n \lambda_n R_n(t)
\]

Since all units are operating at the initial instant \( t = 0 \), we have \( R_n(0) = 1 \) and \( R_k(0) = 0 \) for \( k < n \). For the solution of these equations see Mathematical methods.
of reliability by Gnedenko.

\[ R_0(t) = \frac{n! \Lambda_1 \cdots \Lambda_N}{2 \pi i} \int_{a-i \infty}^{a+i \infty} \frac{e^{zt}}{z(z + \lambda_1) \cdots (z + n \lambda_n)} \frac{dz}{z} \]

Obviously, \( R_0(t) \) is the probability that all units failed prior to the instant \( t \), i.e., the probability that the system failed. Therefore,

\[ R_0(t) = F(t) = 1 - R(t). \]

2.6 BAYESIAN RELIABILITY ANALYSIS:

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

2.6.1 Essential Elements in a Bayesian Reliability Analysis:

There are several elements in a good Bayesian reliability analysis.

i) A detailed justification and analysis of the prior distribution selected, with a clear understanding of the mathematical implications of the prior.

ii) A thorough documentation of the data sources in identifying and selecting the prior.
iii) A preposterior analysis of the prior distribution with hypothetical test results,
iv) A clearly defined posterior distribution on the parameter of interest.
v) An analysis of the sensitivity of the Bayesian inference to the prior model selected.

2.6.2 **Bayesian Estimation Theory**:

Bayes theorem, which was introduced in section (1.2), is the fundamental tool used to arrive at Bayesian inferences. Before proceeding to a discussion, let us introduce the following notations:

\[ g(\theta) : \text{the prior density on } \Theta \]

\[ f(x_1, \ldots, x_n/\theta) : \text{the joint density of a sample of size } n \text{ from } f(x/\theta), \]

\[ f(x_1, \ldots, x_n, \theta) = f(x_1, \ldots, x_n/\theta)g(\theta) : \text{The joint probability distribution of } x_1, \ldots, x_n \text{ and } \Theta \]

\[ f(x_1, \ldots, x_n) = \int_{-\infty}^{\infty} f(x_1, \ldots, x_n, \theta) \, d\theta : \text{The joint marginal density of the sample observations.} \]

\[ g(\theta|x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n, \theta)}{f(x_1, \ldots, x_n)} : \text{the posterior density of } \Theta. \]

Bayes methods in Reliability involve a known prior distribution, and data \((x_1, \ldots, x_n)\) to estimate the particular value of \(H\) drawn. It was also indicated that,
since the prior distribution is known, the posterior distribution is available, as well as estimate based on minimizing the expected loss.

Suppose that $\theta$ is an estimate of $H$ and that the loss function is

$$L(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$$

This loss function is after called the quadratic loss function. The Bayes approach is to select the estimate of $H$ that minimizes the expected loss with respect to the posterior distribution.

Which is clearly minimized when

$$\theta = E(\theta | x_1, \ldots, x_n) = \int_{-\infty}^{\infty} \theta g(\theta | x) \, d\theta$$

and thus the posterior mean $E(\theta | x)$ is a Bayes estimator for $\theta$.

**Absolute Error loss Function** : The absolute error loss function

$$L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$$

assumes that the loss is proportional to the absolute value of the estimation error. The Bayes estimator will be that estimator that minimizes the expected loss. Chernoff and Moses (1959) show that the value of $\theta$ is any medium of the posterior distribution of $\theta$ given $x$.

2.6.3 **Estimation Based on a Sufficient Statistic** :

Let $X$ be a r.V from the conditional probability distribution $f(x | \theta) = \prod_{i=1}^{n} f(x_i | \theta)$. If there exists a set of jointly sufficient statistics $Z' = (z_1, \ldots, z_q)$
for estimating \( \Theta' = (\Theta_1, \ldots, \Theta_p) \), then, according to the Neyman factorization criterion

\[
f(x / \Theta) = f(z / \Theta) h(x)
\]

where \( f(z / \Theta) \) is the conditional probability distribution of \( Z \) given \( \Theta \) and \( h(x) \) is some nonnegative function of \( x \) not involving \( \Theta \).

If \( g(\Theta / x) = g(\Theta / z) \), then the sufficient statistic \( Z \) can be used in lieu of \( x \) in constructing Bayes estimators.

If is easily seen that

\[
g(\Theta / x) = \frac{f(x / \Theta) g(\Theta)}{\int \theta f(x / \Theta) g(\Theta) \, d\Theta} = \frac{f(z / \Theta) h(x) g(\Theta)}{\int \theta f(z / \Theta) h(x) g(\Theta) \, d\Theta}
\]

\[
= \frac{f(z / \Theta) g(\Theta)}{\int f(z / \Theta) g(\Theta) \, d\Theta} = g(\Theta / z)
\]

In particular, for a quadratic loss function, the Bayes estimator for \( \Theta \) becomes,

\[
\hat{\Theta} = E(\Theta / x_1, \ldots, x_n) = E(\Theta / z_1, \ldots, z_q)
\]

2.7 LIFE TESTING:

Reliability estimation methods assume the existence of failure data obtained from life testing. Such testing depending on a random sample of \( n \) devices from a hypothesized population of such devices is placed on test under specified environmental conditions, and failure times of
some or all of the units are observed. If each device that fails is immediately replaced by a new one, the resulting life test is called a test with replacement, otherwise, the life test is said to be a test without replacement.

A time-truncated (or censored) life test is one which is terminated after a fixed period of time has lapsed, whereas an item-censored life test is one which is terminated after a prespecified number of failures have occurred. The time-censored life tests are often referred to as Type I censored life test, whereas item-censored life tests are often called Type II censored life tests.

In order to induce failures of very-high reliability devices, special testing methods known as accelerated life tests are used.

2.7.1 Life test Experiments:

Epstein (1958, 1960) considered several possible life test experiments, some of them are:

i) Testing is terminated after a prespecified number of failures have occurred, failures are replaced (Type II/item-censored testing with replacement).

ii) Testing is terminated after a prespecified number of failures have occurred, failures are not replaced (Type II/item-censored testing without replacement).

iii) Testing is terminated after a prespecified time has elapsed, failures are replaced (Type I/time censored testing with replacement).
Testing is terminated after a prespecified time has elapsed, failures are not replaced (Type I/time truncated testing without replacement).

In case (i) and (ii), the number of failures is a fixed constant and time is the r.V., whereas in case (iii) and (iv) the opposite is true.

As seen above, a life test can be terminated at a particular time, or often a particular number of failures occur, or, in fact, all items can be tested to failure. In test planning, the accuracy of the resulting statistical estimation will be determined by the number of failures obtained. Also, the more items placed on test, the quicker one will obtain a preselected number of failures.

Let $T_1 \leq T_2 \leq \ldots \leq T_r$ denote the r.V.'s corresponding to the observed sample of r ordered failure times. If the number of failures is an r.V., it is denoted by $R$. Let $n$ items are placed on test and let $t_0$ denote the test termination time. Let $R$ represent the total test time accumulated on all items including those that failed and those that did not fail prior to test termination. Corresponding to above four cases we have

1. $T = n \ T_r$, (r specified, $T_r$ random),
2. $T = \sum_{i=1}^{r} T_i + (n-r) \ T_0$, $r \leq n$ (r specified, $T_r$ random),
(iii) \( T = n t_0 \) (to specified, \( R \) random),

\[
\text{(iv)} = \sum_{i=1}^{R} T_i + (n-R)t_0, \ R \leq n \ (\text{to specified, } R \text{ random}).
\]

2.8 **ACCELERATED LIFE TESTS:**

Many devices such as electronic items have very high reliability when operating within their intended normal use environment. This presents problems in measuring the reliability of such devices because a very long period of testing under the actual operating conditions would be required to obtain sufficient data to estimate the reliability. Even if this testing could be accomplished, the time frame is such that the devices may become obsolete before their reliability is established due to the high rate of technological advances. Also, it would be difficult to conduct the testing in a laboratory.

One solution to the problem of obtaining meaningful life test data for high reliability devices is accelerated life testing. This type of testing involves observing the performance of these kinds of devices operating at higher stress levels than usual to obtain failures more quickly. In order to shorten product life, it is a well-established engineering practice to use certain stresses or accelerating variables, such as higher levels of temperature, voltage, pressure, vibration, etc., than the normal operating level.
The main difficulty of accelerated life testing lies in using the failure data obtained at the accelerated, or higher stress, conditions to predict the reliability, mean life, or other quantities under the normal use condition. Extrapolation from the accelerated stresses to the normal use stress is done by choosing an appropriate model, called an acceleration model. The choice of an acceleration model calls for a knowledge of the variation of failure behaviour with environment. In parametric method, this involves functional relationship between the parameters of the failure distributions and the environmental stresses. The relationship may also involve unknown parameters. In nonparametric approaches, where no specific form of the failure distribution is specified, the change in the failure distribution due to a change in environmental stress is assumed. In either the parametric or nonparametric, all unknown parameters must be estimated from the accelerated test data in order to extrapolate to the normal use stress.

Four acceleration models are used, i.e. power rule model, the Arrhenius model, the Eyring model, and the generalized eyring model. These models will be discussed by Mann, Schafer, and Singpurwall (1974).

2.8.1 Acceleration Models:

The use of accelerated life testing to make inferences about the normal use life distribution requires a model
to relate the life length to the stress levels that
are to be applied to the items being tested. This model
is referred to as the acceleration model.

Here some acceleration models that has been used
in parametric and nonparametric method will be described
briefly.

In parametric, suppose the life time random variable
$X_1^0$ of items in an environment described by a constant
stress level $V_1$ has a probability distribution $F^0(t; \theta_1)$
depending on a vector of parameter $\theta_1$. Two assumptions
which are made (Mann, Schafer, and Singpurwalla, 1974)
are
i) The change in stress level does not change the type
of the life time distribution $F^0(t; \theta)$, but changes only
the parameter values.

ii) The relationship between the stress level $V$ and the
parameters $\theta$, say $\theta = m(V, \alpha, \beta...)$, is known except for
one or more of the acceleration parameter $\alpha, \beta...$, and
that the relationship is valid for a certain range of
the elements of $V$. The objective here is to obtain
estimates of the parameters $\alpha, \beta...$ based on life test
data obtained at large values of $V$ and make inferences
about $\theta$ for the normal use stress $V_0$.

The exponential distribution with parameter $\lambda$ is
widely used as a lifetime distribution. So the acceleration
models will be discussed here for exponential distributions. Several authors have considered other lifetime distributions such as weibull (Mann, 1972, and Nelsen, 1975), extreme value (Meeker and Nelson, 1975, and Nelson and Meeker, 1978), and lognormal (Nelson and Kielpinski, 1976), for example. Suppose that under constant application of single stress at level \( V_1 \), the item being tested has an exponential lifetime distribution with mean \( \mu_1 \) given by

\[
f^0(t; \lambda_1) = \lambda_1 e^{-\lambda_1 t}, \quad t \geq 0, \lambda_1 > 0
\]

Then \( \mu_1 = 1/\lambda_1 \) is the mean time to failure under stress level \( V_1 \). The following acceleration models (relationships between \( \lambda_1 \) and \( V_1 \)) have been suggested in the literature.

i) The Power Rule (or inverse power) Model:

This model can be derived by considerations of kinetic theory and activation energy. This model has applications to fatigue testing of metals, the dielectric breakdown of capacitors, and aging of multicomponent systems. The model is

\[
\mu_1 = \alpha V_1^{-\beta}, \quad \alpha > 0, \beta > 0
\]

and this implies that the mean time of failure \( \mu_1 \), decreases as the \( \beta \)th power of the applied voltage \( V \). It is desirable to estimate \( \alpha \) and \( \beta \) from life test data at stress levels \( V_1, \ldots, V_k \) and make inferences about \( \mu_0 = 1/\lambda_0 \) at the normal use stress \( V_0 \).
11) **The Arrhenius Model**: This model expresses the degradation rate of a parameter of the device as a function of its operating temperature. It is usually applied to thermal aging and is applicable to semiconductor materials. Here
\[ \lambda_1 = \exp(\alpha - \beta/V_1) \]
is the model, where \( V_1 \) denotes the temperature stress and \( \alpha \) and \( \beta \) are unknown parameters to be estimated in order to make inferences about \( \lambda_0 \) at normal temperature level \( V_0 \).

11) **The Eyring Model for a Single Stress**: This model can be derived from principles of quantum mechanics and expresses the time rate of degradation of some device parameter as a function of the operating temperature. Here
\[ \lambda_1 = V_1 \exp(\alpha - \beta/V_1) \]
is the model.

iv) **The Generalized Eyring Model**: This model has application to accelerated testing of devices subjected to a constant application of two types of stresses, one thermal and one nonthermal. The model is
\[ \lambda_1 = \alpha T_1 \exp(-\beta/K T_1) \exp(\gamma V_1 + \delta V_1/K T_1) \]
where \( \alpha, \beta, \gamma \) and \( \delta \) are unknown parameters to be estimated, \( K \) denotes Boltzmann's constant, whose value is \( 1.38 \times 10^{-16} \) erg/degree Kelvin, and \( T_1 \) is thermal stress level and \( V_1 \) is the nonthermal stress. In the absence of a nonthermal stress, this model reduces to
\[ \lambda_1 = \alpha T_1 \exp \left( -\beta / T_1 \right) \]

Chernoff (1962) considered an acceleration model for exponential life time with mean \( \mu_1 = (\alpha V_1 + \beta V_1^2)^{-1} \) where \( \alpha > 0 \) and \( \beta > 0 \) were unknown parameters. Thus \( \lambda_1 \) was a quadratic function of the stress level. Chernoff also considered models for three dimensional vector stresses

\[ V_1 = (V_{11}, V_{21}, V_{31}) \]

In partially nonparametric approaches to inference from accelerated life tests, no particular form of the life time distribution is assumed, but an acceleration model is used (see, for example, Shaked, Zimmer, and Ball, 1979; Selhuraman and Singpurwalla, 1982; Shaked and Singpurwalla, 1982; Basu and Ebrahimi, 1982; and Shaked and Singpurwalla, 1983). Shaked, Zimmer, and Ball (1979) assumed that the \( K \) accelerated stress levels \( V_1, \ldots, V_K \) were selected of stresses \( V_1, V_j, 1, j = 0, 1, \ldots, K \), a known function \( m \) existed. Therefore the life time distributions satisfied.

\[ F_{V_j}(t) = F_{V_1}[m(\alpha, V_j, V_1, t)], t > 0, \]

where \( \alpha \) is an unknown parameter, the form of \( F_{V_1} \) is not assumed to be known. Various choices of \( m \) gives the power rule, Arrhenius, Eyring, etc., acceleration models. The other references assume models for special cases of \( m \).
In a totally nonparametric setting, there is no assumption made about the form of the lifetime distribution at the various stress levels nor about the forms of an acceleration model. In this setting, the life distributions are stochastically ordered with respect to increasing levels of stress (Barlow and Scheuer, 1971) or that the lifetime distribution at two distinct stress levels differ only by a scale change. For these procedures, it must be assumed that failure data are available from the normal use stress as well as from accelerated stresses.

The design aspects of accelerated life testing experiments involve the selection of stress levels. The number of stress levels, and the number of items to be tested at each stress level. A null-designed estimaters and allow for censoring.

****
CHAPTER - III

BAYESIAN INFERENCE IN RELIABILITY AND ACCELERATED LIFE TESTING

3.1 HISTORY ON BAYESIAN APPROACH TO RELIABILITY :

3.1.1 Introduction:

In the Bayes approach, the parameter to be estimated. Say \( \theta \) in the failure model \( f(t; \theta) \), is presumed to have associated with it a weighting function, \( g(\theta) \), which may or may not represent a probability density. If it does not, it is called an improper prior. The use of such a prior function follows naturally either from the assumption that \( \theta \) is, in fact, a random variable or from the desire to more efficiently use data already available. For more details on introduction to Bayesian Reliability are in chapter IIInd.

3.1.2 Bayesian approach to Reliability:

Canfield, R.V. (1970), in a paper dealing with the estimation of reliability in the exponential failure model, uses the assumption that over estimation of \( R(t) \) is more than underestimation to justify a loss function of the form

\[
L(\hat{R}, R) = \begin{cases} 
K_1 \left( \frac{\hat{R}}{R} - 1 \right)^2 & \text{if } \hat{R} < R \\
K_1 \left( \frac{\hat{R}}{R} - 1 \right)^2 + K_2 \left( \frac{\hat{R}}{R} - 1 \right) & \text{if } \hat{R} > R 
\end{cases} \quad \ldots 3.1.2.1
\]

He uses a beta prior
which is a natural conjugate of the exponential model
(the posterior density is also beta). He concludes that
when no prior information is available \( p = q = 1 \) implying
g is uniform) and a symmetric loss is used \( K_2 = 0 \),
then the resulting estimator is the minimum variance
unbiased estimator of the reliability.

\[
R(t) = e^{-\alpha t}
\]

Springer and Thompson (1966), consider a series
system of components or subsystems, each of which exhibits
a binomial failure model, and for each of which the
prior density is of the beta form (the natural conjugate
to the binomial). They construct Bayesian confidence
bounds for the product of the binomial parameters( the
product of the reliabilities is the reliability of the
system) by deriving the posterior density of the product
using Mellin transform techniques. Papadopoulos and
Tsokos (1974), derived estimates for the reliability
of a binomial failure model when the prior density is
either uniform or beta. In addition, the first two
moments of the estimators were given and the use of gamma
theory techniques for estimation were investigated.

Pugh (1963), examined the applicability of Bayesian
methods to the development of confidence bounds for the
reliability of an exponential failure model with uniform

\[
g(R ; p, q) = \frac{1}{\beta(p,q)} R^{p-1} (1-R)^{q-1}, R \in (0,1)
\]

...3.1.2.3
prior. For a life test until K failures and a mission
time \( t_m \) he obtains the result

\[
P[R > R_o] = 1 - \sum_{\gamma = 0}^{k} \frac{R(t_k/t_m)^{\gamma+1}}{\gamma!} \left[-\log R(t_k/t_m)+1\right]^{\gamma}
\]

EL-Sayyad (1967), considered a large number of estimators
for the parameter of an exponential density assuming
a prior of the form

\[
g(\theta) = \theta^{a-1} e^{-b\theta}, \quad 0 < \theta < \infty, \quad a, b > 0
\]

If the loss function is of the form

\[
L(\hat{\theta}, \theta) = \theta^a (\hat{\theta}^\beta - \theta^\beta)^2
\]

then the estimate of \( \theta \) is

\[
\hat{\theta} = \frac{1}{b+s} \left[ \frac{\Gamma(a+n+\alpha)}{\Gamma(a+n)} \right]^{1/\beta}
\]

where \( S = \sum t_k \) is the sum of the failure times from a
sample of size \( n \). If, however, the loss function is

\[
L(\hat{\theta}, \theta) = (\log \hat{\theta} - \log \theta)^2
\]

then,

\[
\hat{\theta} = \frac{\phi(n)}{b+s}
\]

where

\[
\phi(n) = \exp \left[ \frac{d}{dn} \log \Gamma(n) \right] \xrightarrow{n \to \infty} (\log n - n/2)
\]

Bhattacharya (1967), made a Bayesian analysis of
the exponential model for data subject to either type
of censoring. He considered the uniform, inverted gamma,
It can be considered as the limit of the quotient of the mean life times under stress levels \( s + \Delta s \) and \( s \) respectively.

Many authors has implicitly assumed \( \delta(s) \) to be of the form
\[
\delta(s) = \delta \cdot h(s), \quad h(s) \text{ known} \quad \text{3.2.6.12}
\]
The quantity \( \delta \) unknown. In this case \( \delta \) is related to the relative acceleration constant \( \alpha(s_1,s_2) \) by the equation,
\[
\delta = \frac{\log \alpha(s_1,s_2)}{H(s_2) - H(s_1)} \quad \text{3.2.6.13}
\]
where \( H(.) \) is the integral of \( h(.) \).

Now the basic quantity for a Bayesian inference is the relative acceleration coefficient
\[
\alpha = \alpha(s_1,s_2) \quad \text{3.2.6.14}
\]
The relative acceleration coefficient is unknown and described by a stochastic quantity \( \tilde{\alpha} \) distributed with density \( \pi(\alpha) \).

Let the life time distributions on stress levels \( S_1 \) and \( S_2 \) have densities \( f(\cdot | s_1) \) and \( f(\cdot | s_2) \) respectively. For two independent samples
\[
T_{1,i} \text{ under stress } s_1, \quad T_{2,i} \text{ under stress } s_2, \quad i = 1, \ldots, n
\]
the following statistics are independent:
\[
U_i = \frac{T_{1,i}}{T_{2,i}}, \quad i = 1, \ldots, n \quad \text{3.2.6.15}
\]
and exponential densities as prior for $\theta$. It was also observed that if the improper prior, $g(\theta) = 1/\theta^2$, is used, then the resulting Bayes estimate coincides with minimum variance unbiased estimate (MVUE) of Epstein and Sobel (1953). Finally, a proof was given showing that this is the only prior for which this is true.

Springer and Thompson in (1966, 1967, and 1968) have developed Bayesian confidence bounds for series and parallel systems of components with exponential failure times and for series systems some of whose components are known to have exponential failure rates and for some of which no failure rate information is known.

They use a prior of the form

$$g(R) = \frac{(\beta_0 + 1)\rho_{0}^{+1}}{\Gamma(\rho_{0} + 1)} R^{\beta_0} (\log 1/R)\rho_0 , \quad R = e^{-t_m/\theta}$$

and utilizing Mellin transform techniques to derive the posterior density of the system reliability

$$R_s = \prod_{i=1}^{n} R_i$$

or

$$R_p = 1 - \prod_{i=1}^{n} (1 - R_i)$$

Basically, their results involve a fast way of writing down the partial fraction expansion for the Mellin transform of the system posterior density of $R$. Ferlig (1972), attempted to determine the optimal component priors for the above priors with optimal defined in the sense that
the resulting confidence intervals are as close as possible to the uniformly most accurate (UMA) unbiased confidence intervals. He concluded that the prior used by Springer and Thompson was the best when prior data was available, but that no prior was optimal when no prior information existed ($\beta_0 = r_0 = 0$). Mann (1967), used Montecarlo computer techniques to obtain the values of $r_0$ and $\beta_0$ for each component so that the resulting system prior yielded Bayesian confidence intervals that were optimal in the same sense. Results similar to Fertig's were obtained showing that no choice exists, independent of previous data, which results in optimal confidence intervals.

EL-Sayyad (1967), applied a method of sequential testing proposed by Lindley (1961), to the exponential model. The method begins with a measure of information suggested by Shannen which estimates the amount of knowledge about $\theta$ which we have in $g(\theta)$:

$$I = \int g(\theta) \log g(\theta) d\theta \quad \text{3.1.2.12}$$

when additional data $X = (x_1, x_2, \ldots, x_n)$ has been obtained we have a new amount of information,

$$I_n = \int h(\theta|x) \log h(\theta|x) \, d\theta \quad \text{3.1.2.13}$$

which uses the posterior density. We may then treat the posterior as prior and obtain more information. The method says, in effect, collect information until $I_n$
reaches a certain value. Furthermore, the method is generalizable by choosing a monotone function \( \hat{\Theta} (\Theta) \) of \( \Theta \) and defining

\[
I_{\hat{\Theta}} = I_\Theta + \int g(\Theta) \log \frac{d\hat{\Theta}}{d\Theta} d\Theta \quad \ldots \ldots 3.1.2.14
\]

clearly, the choice of \( \hat{\Theta} \) will have a strong effect on the form of the posterior from which we finally make a Bayesian inference.

El-Sayyad uses

\[
g(\Theta) = \frac{ab}{\Gamma(a)} \Theta^{a-1} e^{-b\Theta}, \quad 0 < \Theta < \infty, \quad a, b > 0
\]

for the prior of \( \Theta \), form which

\[
I_\Theta = a \log b - \log \Gamma(a) - a + (a-1) [\hat{\Theta}(a) - \log b]
= 1/2 \log \frac{b^2}{2\pi a} - \frac{a + 1}{2a} \quad \ldots \ldots 3.1.2.15
\]

The \( \hat{\Theta} \) here is the derivative of \( \log \Gamma(a) \), as before.

He then considers several forms of \( \hat{\Theta}(\Theta) \):

\( \hat{\Theta}(\Theta) = \Theta^c \), for which

\[
I_{\hat{\Theta}} = 1/2 \log b + \log 2 - \log \sqrt{2\pi} - 1/2 + 1/4a, \quad \ldots \ldots 3.1.2.16
\]

\( \hat{\Theta}(\Theta) = \log \Theta \), for which

\[
I_{\hat{\Theta}} = 1/2 \log \frac{a}{2\pi} - 1/2, \quad \ldots \ldots 3.1.2.17
\]

\( \hat{\Theta}(\Theta) = 1/\Theta \), for which

\[
I_{\hat{\Theta}} = I_\Theta + 2 \text{E}(\log \Theta)
= 1/2 \log \left( \frac{a}{2\pi b^2} \right) - \frac{(a+1)a}{2} \quad \ldots \ldots 3.1.2.18
\]
Deely, Tierney, and Zimmer (1970), have considered the use of the 'Maximum Entropy Principle' (MEP) in choosing a prior for the binomial and exponential models. The principle defines the optimal choice to be that prior which, within the constraints of available data, maximizes the entropy of the prior density of

\[ S_\theta = - I_\theta , \]  

that is, which minimizes the presumed measure of information available about \( \theta \). They introduce the notion of a "least favourable distribution". If \( G \) is a family of prior distributions, then \( G_0 \) is least favourable over \( G \) if its minimum expected loss (the loss expected for the Bayes estimator) is greater than that of any other member of \( G \). Under the assumption of a quadratic loss, it is shown that a least favourable distribution may be a better choice of prior than the MEP choice.

The difficulties arising from the MEP were approached in a different manner by Jaynes (1968), who suggested that the principle could be generalized by consideration of the measure.

\[ S_\Phi = - \int_\Omega g(\theta) \log \left[ \frac{\Phi(\theta)}{\phi(\theta)} \right] d\theta \]  

where \( \Phi \) is, as El-Sayyad (1966), noted, an appropriate monotone function of \( \theta \). He further suggested that the choice of \( g(\theta) \) should be made using transformation group techniques. Essentially, such an approach says to choose
a g in such a way that changes in the parameters of g(θ) do not change the entropy measure. For example, if θ_1, is a location parameter and θ_2 is a scale parameter, then we seek a prior g(θ_1, θ_2) such that

\[ g(θ_1, θ_2) = a g (θ_1 + 8, aθ_2) \]

The solution is of the form

\[ g (θ_1, θ_2) = \frac{\text{Constant}}{θ_2} \quad \ldots \quad 3.1.2.21 \]

In a similar way, it is arranged that the binomial model leads to the equation

\[ θ (1 - θ) g′(θ) = (2θ - 1) g(θ), \]

whose solution is

\[ g(θ) = \frac{\text{Constant}}{θ(1-θ)} \quad \ldots \quad 3.1.2.22 \]

Harris and Singpurwalla (1968) dealt with the application of Bayesian estimation methods to failure models with hazard rates of the form

\[ h(t) = θ α t^{α-1}, \quad 0 < t < \infty, \quad θ, α > 0 \]

when complete samples are available. This hazard rate corresponds to the Weibull model and becomes the exponential when α = 1. They derived estimators for θ in the exponential and Weibull cases with the uniform, two-point, and gamma priors. They also investigated the estimation of α where θ is known but obtained a solution only for the two point prior for α. Soland (1968) determined a Bayesian acceptance sampling procedure for θ when α is known for the Weibull
model and data subjected to Type II censoring (m failures out of n items tested simultaneously), using a gamma prior. Later, Soland (1969), extended the results to the case in which both \( \alpha \) and \( \theta \) are unknown.

Canavos and Tsokos (1973), constructed a Bayesian analyses of the scale and shape parameters and the reliability function of the weibull model for the uniform, exponential, and inverted gamma prior densities. The results are a direct generalization of the work of Bhattacharya (1967) in which the one parameter exponential model was concerned. The two parameter case was analyzed under the assumption that the parameters were independently distributed. The estimates and their variance for Type II censored data were obtained for the parameters and the reliability function, \( R(t) = e^{-t^{\alpha}/\theta} \), and were compared with the corresponding minimum variance unbiased or maximum likelihood estimator by an analysis of the relative mean squared errors (MSE). The Bayesian estimates were found to have uniformly smaller MSE. Tsokos (1973), continued this work by studying the effect of the choice of a wrong prior on the MSE of the estimators by use of Monte Carlo simulation.

Papadopoulos and Tsokos (1976) have extended the work of Soland (1968) by deriving a Bayes reliability estimator for the case in which both the scale and shape parameters are unknown. The result was compared with the MLE for the reliability using numerical computer simulation techniques.
3.2 BAYESIAN INFERENCE IN ACCELERATED LIFE TESTING:

3.2.1 Accelerated Life Testing:

In many practical important situations it is impossible to make life tests for devices under usual environmental conditions. Especially for components or systems with high reliability it is seldom possible to make life tests under usual environmental conditions (usual stress $S_u$) because testing time would exceed the available time. Therefore a stress $S > S_u$ is applied which reduces the life time. Life tests of this kind are called accelerated life tests. Details on accelerated life testing (ALT) are given in chapter II and also on history is given in chapter IV. More details are given in the monograph 'Statistical Methods in Accelerated Life Testing' by Viertl (1987).

The stochastic description of accelerated life testing is the following. Let $S$ be a one or higher dimensional stress applied to a device. The cumulative distribution function (c.d.f.) of the life time of the device depending on the stress $S$ is denoted by

$$F(t/S) \text{ for } t > 0$$

\[\text{...3.2.1.1}\]

The main objective of ALT is to make inference on

$$F(t/S_u) \text{ for } t > 0$$

\[\text{...3.2.1.2}\]

the cdf of the life time under usual stress $S_u$ from observations of the life time under accelerating stress
levels $S > S_u$, Here $>$ means for the corresponding cdf's

$$F(t/S) > F(t/S_u) \text{ for all } t > 0.$$  

For two stress levels $S_1 < S_2$ the relation between $F(t/S_1)$ and $F(t/S_2)$ can be described by a so-called acceleration function $a(t; S_1, S_2)$, $t > 0$:

$$F(t/S_2) = F(a(t; S_1, S_2)/S_1) \text{ for all } t > 0 \ldots 3.2.1.3$$

It was assumed that the relation between stress $S$ and corresponding life time distribution to be given by the dependence of a statistical parameter $\Theta$ on the applied stress $S$, i.e.

$$F(t/\Theta(S)) \text{ for } t > 0 \ldots 3.2.1.4$$

Also it was assumed that the acceleration function is linear, i.e.

$$a(t; S_1, S_2) = \alpha(S_1, S_2).t \text{ for all } t > 0 \ldots 3.2.1.5$$

In this case the problem is to estimate $\alpha(S_u, S)$ and the cdf $F(t/S_u)$ is related to the cdf $F(t/S)$ for $S > S_u$ by

$$F(t/S_u) = F\left(\frac{1}{\alpha(S_1, S_2)}.t/S\right) \text{ for all } t > 0 \ldots 3.2.1.6$$

Another approach is to use the failure rate $r(t) = f(t) / \bar{F}(t)$, with reliability function $\bar{F}(t) = 1 - F(t)$ for all $t > 0$, for the stress life time relationship. Here the following equality is used

$$F(t) = 1 - \exp \left[ - \int_0^t r(x) \, dx \right] \ldots 3.2.1.7$$

The stress dependence of the life time distribution is
modelled by a stress dependent failure rate
\[ r(t/S) = \lambda(S) r_0(t) \text{ for all } t \geq 0 \]
with \( r_0(.) \) known function and different assumptions on \( \lambda(s) \). This method is described in section 4.

3.2.2 Bayesian Inference in Life Testing and Reliability

Estimation:

The idea of Bayesian inference is to model all unknown quantities by stochastic quantities (also called random quantities or random variables) and to describe the uncertainty about these quantities by probability distribution.

Let \( X \) be the life time of a device with stochastic model \( f(t/\Theta) \) depending on the parameter \( \Theta \), and let it is described by a stochastic quantity \( \Theta \) with a-priori distribution \( \pi(\Theta) \). If life data are observed the information in the data \( D, D = (t_1, \ldots, t_n) \) is used via Bayes' theorem to obtain the a-posteriori description of the uncertainty about \( \Theta \). This a-posteriori description is the a-posteriori distribution \( \pi(\Theta/D) \) of \( \Theta \) obtained by

\[ \pi(\Theta/D) \propto \pi(\Theta) \cdot L(\Theta; D) \]

where \( L(\Theta,D) \) is the likelihood function depending on the stopping rule of the sampling procedure and \( \propto \) stands for proportional (upto a constant). In extensive form Bayes' theorem can be written in the following way.
\[ \pi(\theta/t_1, \ldots, t_n) = \frac{L(\theta;t_1, \ldots, t_n)\pi(\theta)}{\int_{\Theta} L(\theta; t_1, \ldots, t_n)\pi(\theta) \, d\Theta} \] ...3.2.2.2

Here \( \Theta \) is the set of all possible values of the parameter \( \theta \), called parameter space.

For reliability analysis the predictive density of the life time after the observation of the data \( D \) is given by

\[ f(t/D) = \int_{\Theta} f(t/\theta) \pi(\theta/D) \, d\theta \] ...3.2.2.3

The predictive reliability function \( F(t/D) = 1 - F(t/D) \) for \( t \geq 0 \) is obtained by

\[ F(t/D) = \int_{t}^{\infty} \int_{\Theta} f(x/\theta) \pi(\theta/D) \, d\theta \, dx \] ...3.2.2.4

To obtain interval estimates of the parameter \( \theta \), Bayesian highest a-posteriori density regions (HPD-regions) are obtained using the a-posteriori density \( \pi(\theta/D) \) for \( \theta \).

For confidence level \( 1 - \alpha \) the HPD-region for \( \theta \) is defined by

\[ \Theta^* = \{ \theta : \pi(\theta/D) \geq C_{1-\alpha} \} \] ...3.2.2.5

where \( C_{1-\alpha} \) is the largest constant such that \( P_{\Theta}[\theta \in \Theta^*] = 1 - \alpha \)

3.2.3 General Parametric Inference:

Let the life time distribution function \( F(t/S) \) be continuous with density \( f(t/\theta(s)) \) and stress dependent parameter \( \theta \), i.e.

\[ \theta(s) = \Psi(S, C_1, \ldots, C_k) \] ...3.2.3.1
where \( \psi(.) \) is known and the quantities \( C_1, \ldots, C_k \) are unknown. The vector \( C = (C_1, \ldots, C_k) \) is described by a stochastic quantity.

\( \tilde{C} \) with a priori distribution \( \pi(C) \).

Now Bayesian analysis is carried out in the following way. If lifetime data \( D = (t_{ij} : i = 1, \ldots, m; j = 1, \ldots n_i) \) are observed for stress combination \( S_1, \ldots, S_m \) the \( \text{a-posteriori} \) distribution \( \pi(C|D) \) of \( \tilde{C} \) is obtained via Bayes' theorem.

\[
\pi(C|D) \propto \pi(C) L(C;D) \quad \ldots \quad 3.2.3.2
\]

where \( L(C;D) = L(C;t_{ij}) \) is the likelihood function which is in the simplest case of uncensored data given by

\[
L(C;t_{ij}) = \prod_{i=1}^{m} \prod_{j=1}^{n_i} f(t_{ij}/C,S_i) \quad \ldots \quad 3.2.3.3
\]

The likelihood for a general sampling plan is obtained from a theorem by Barlow and Prochan (1980) in the following way:

Let \( r(t) = f(t)/F(t) \) be the failure rate.

Then

\[
f(t) = r(t) e^{-\int_0^t r(u) du} \quad \ldots \quad 3.2.3.4
\]

using the cumulated hazard function

\[
R(t) = \int_0^t r(u) du \quad \ldots \quad 3.2.3.5
\]

the likelihood function on stress level \( S_1 \) is in the simplest case.
For a general sampling plan with failure times $t_1, \ldots, t_k$ and withdrawals $w_1, \ldots, w_{m_1}$ using the number $n(t)$ of functioning units on test under stress $S_1$ as a function of time $t$ the likelihood conditional given data

$$D = (t_1, \ldots, t_k, w_1, \ldots, w_{m_1})$$

$$L[r(u); u > 0/D] = \prod_{j=1}^{n_1} r(t_{ij}) e^{-\int_{0}^{\infty} n(u) r(u) du} \ldots 3.2.3.7$$

where the first product is equal to 1 for $k_1 = 0$.

The a-posteriori distribution $\pi(C/D)$ of the common characteristic quantity $C$ can be used in different ways.

1) **Predictive Density under Design stress**

By the model $\Theta(s) = \Psi(S, C)$ for the stress dependence of the stochastic parameter $\Theta$ we obtain for the parameter $\Theta_u = \Theta(S_u)$ under design (usual) stress $S_u$

$$\Theta_u = \Theta(S_u) = \Psi(S_u, C) \ldots 3.2.3.8$$

From the a-posteriori distribution $\pi(C/D)$ of $\tilde{C}$ the distribution $\pi(\Theta_u/D)$ of $\tilde{\Theta}_u$ can be calculated. This distribution is used to calculate the predictive density of the life time under design stress $S_u$ in the following was:

$$f(t/S_u, D) = \int_{\Theta_u} f(t/\Theta_u) \pi(\Theta_u/D) d\Theta_u \ldots 3.2.3.9$$
ii) **Estimation of the Distribution Function under Design Stress:**

By \( F(t/S_y) = F(t/\psi(S_y, C)) \) and the description of the unknown quantity \( C \) by a stochastic quantity \( \tilde{C} \), a stochastic description of \( F(t/S_u) \) is given by

\[
\tilde{F}(t/S_u) = F(t/\psi(S_u, \tilde{C})) \quad \text{for all } t \geq 0
\]

Using the a-posteriori distribution \( n(C/D) \) of \( \tilde{C} \) after observation of the data \( D \), we obtain the stochastic quantity

\[
\tilde{F}(t/S_u, D)
\]

which can be used to obtain point and interval estimates in a Bayesian manner.

(a) **Point Estimates for \( F(t/S_u) \):**

Point estimations of the cdf \( F(t/S_u) \) of the life time are possible on different levels of using the information in the model and the data.

**First** the a-posteriori estimate

\[
\hat{C} = E_{n(C/D)} \tilde{C}
\]

...3.2.3.11

can be used to estimate the cdf \( F(t/S_u) \) by

\[
\hat{F}(t/S_u) = F(t/\psi(S_u, \hat{C})) \quad \text{for } t > 0 \quad ...3.2.3.12
\]

**Second** using the expectation of \( \tilde{\theta}(S_u) = \psi(S, \tilde{C}) \) and estimating the c.d.f by

\[
\hat{F}(t/S_u) = F(t/E_{n(C/D)} \psi(S_u, \tilde{C})) \quad \text{for } t > 0 \quad ...3.2.3.13
\]

**Third** the c.d.f of the life time under design stress can be estimated pointwise by
The integrations for the second and third estimates can be complicated and time consuming. Therefore in section 5 the concept of semi-sufficiency introduced by Willing (1987) is described.

(b) Interval Estimation for $F(t/S_u)$:

The distribution of the stochastic process

$$\hat{F}(t/S_u, D) = F(t/\Psi(S_u, C)),$$  \( t \geq 0 \) \( \ldots 3.2.3.15 \)

can be derived from the distribution $\pi(C/D)$ and HPD-regions for $F(t/S_u, D)$ can be constructed to obtain band estimates for the c.d.f of the lifetime under design stress $S_u$.

3.2.4 Time Transformed Exponential Model:

The time transformed exponential model introduced by R. Barlow and A. Wu is assumed to be of the form

$$F(t/S) = 1 - e^{-\lambda(s)R_0(t)}$$ \( \ldots 3.2.4.1 \)

where $R_0(.)$ is known and $\lambda(s)$ is a function of the stress $S$. If the function $\lambda(s)$ is assumed to be of the form

$$\lambda(s) = \Psi(S, C_1, \ldots, C_k)$$ \( \ldots 3.2.4.2 \)

similar to equation (3.2.3.1) for the dependence of the parameter $\Theta$ on the stress $S$ in section (3.2.3) the analysis is analogous to the analysis in section (3.2.3.).

If the function $\lambda(s)$ is of the form

$$\lambda(s) = \lambda \Psi(s) \text{ with } \Psi(s) \text{ known}$$ \( \ldots 3.2.4.2 \)
the analysis is comparatively easy. For a-priori distribution $\pi(\lambda)$ of the stochastic quantity $\tilde{\lambda}$ describing $\lambda$ and data $D$

$$D = (t_{ij}) \quad i = 1, \ldots, m$$

$$j = 1, \ldots, n_i$$

obtained under $m$ stress levels $S_1, \ldots, S_m$ for uncensored data $D$ the likelihood is given by

$$L(\lambda/D) = \left( \frac{\lambda^{dR_o(t)}}{dt} \right) \prod_{i=1}^{m} \prod_{1}^{n_i} \left[ \psi(S_i) e^{-\lambda \psi(S_i) R_0(t)} \right]^{n_i}$$

...3.2.4.4

The a-posteriori distribution $\pi(\lambda/D)$ is obtained by Bayes' theorem.

The predictive density of the lifetime under design stress $S_u$ is very more easier to compute than in the general parametric model, i.e.

$$f(t/S_u, D) = \int f(t/\lambda)\pi(\lambda/D)d\lambda \quad \ldots3.2.4.5$$

where $\Lambda$ denotes the set of all possible values of $\lambda$.

The distribution of $F(t/S_u) = F(t/\lambda(S_u)) = F(t/\lambda \cdot \Psi(s))$ is calculated in an analogous way as in the general parametric model (compare section (3.2.3)).

3.2.5 Semi-Sufficiency and Computations:

The computations in section (3.2.3) can be done analytically only for exponential families. Exponential families are two restrictive which is seen by the Weibull distribution. In order to carry out necessary integrations for higher dimensions of the parameter $\theta$ the concept of
semi-sufficiency is useful to reduce the dimensions of the integrations.

Let $X$ be a stochastic quantity with density $f(x/\theta)$ and $\theta = (\theta_1, \theta_2)$ is decomposed into two subvectors $\theta_1$ and $\theta_2$. A statistic $t(x/\theta_2)$ is called semi-sufficient for $\theta$ if the a-posteriori distribution of $\theta_1$ given $\theta_2 = \theta_2$ depends on $X$ only through $t(x/\theta_2)$ for all a-priori distributions of $\theta$.

$$\pi(\theta_1/\theta_2, x) = g(\theta_1/\theta_2, t(x/\theta_2))$$ ...3.2.5.1

By a theorem proved by willing (1987) a statistic $t(x/\theta_2)$ is semi-sufficient for $\theta$ if and only if the conditional density $f(x/\theta)$ factors in the following way

$$f(x/\theta) = h(t(x/\theta_2)/\theta_1, \theta_2) \psi(x/\theta_2)$$ ...3.2.5.2

where $h(.)$ is depending on $x$ only through $t(x/\theta_2)$ and $\psi(.)$ doesn't depend on $\theta_1$.

The concept of semi-sufficiency is useful for so-called semi-exponential families of densities. Let $X$ be a stochastic quantity whose distribution $f(x/\theta)$ is depending on a parameter vectors $\theta = (\theta_1, \theta_2)$ where $\theta_1$ and $\theta_2$ are subvectors for which the presentation (3.2.5.1) is valid. If the stochastic model $f(x/\theta_1, \theta_2)$ belongs to an exponential family in $\theta_1$ for given $\theta_2$, then $t(x/\theta_2)$ is a semi-sufficient statistic of a fixed dimension. This means the dimension of the statistic doesn't increase when taking a sample of $X$.

A family of distributions of a stochastic quantity $X$ is called a semi-exponential family of dimension $K$ with
parameter vector \( \Theta \) if and only if \( \Theta \) can be decomposed in to \( \Theta = (\Theta_1, \Theta_2) \) such that \( f(x/\Theta) \) has the following presentation

\[
f(x/\Theta_1, \Theta_2) = G_1(\Theta_1/\Theta_2) \cdot G_2(\Theta_2) \cdot H(x/\Theta_2).
\]

\[
\exp \left[ - \sum_{j=1}^{k} \psi_j(\Theta_1/\Theta_2) \cdot T_j(x/\Theta_2) \right] \quad \ldots \text{3.2.5.3}
\]

For a sample of a distribution belonging to a semi-exponential family with presentation (3.2.5.3) a conjugate family of a-priori distributions for the parameter \( \Theta = (\Theta_1, \Theta_2) \) is given in willing (1987) using hyperparameters \( \alpha_0, \alpha_1, \ldots, \alpha_k, P \) where \( p = p(\Theta_2) \) can be marginal density for \( \Theta_2 \). The hyperparameters are updated by the information from data \( D \) to obtain the a-posteriori distribution

\[
\pi(\Theta_1, \Theta_2/D) = \pi(\Theta_1, \Theta_2/\tilde{\alpha}_0, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_k, \tilde{p}) \quad \ldots \text{3.2.5.4}
\]

For details see willing (1987)

1) **Weibull Distribution**:

If \( X \sim \text{Wei}(\tau, \eta) \) then a sample \( X = x_1, \ldots, x_n \) of \( X \) has joint density

\[
f(x_1, \ldots, x_n/\tau, \eta) = \tau^n \eta^n \left( \prod_{i=1}^{n} x_i \right)^{n-1} \cdot e^{-\tau \sum_{i=1}^{n} x_i^\eta} \quad \ldots \text{3.2.5.5}
\]

Taking \( \Theta = (\Theta_1, \Theta_2) = (\tau, \eta) \) we obtain the factorization (3.2.5.2) with the functions

\[
h(t(x/\Theta_2)/\Theta_1, \Theta_2) = \tau^n \cdot e^{-\tau \sum_{i=1}^{n} x_i^\eta} \quad \ldots \text{3.2.5.6}
\]
The semi-sufficient statistic is of the form
\[ t(x/\theta_2) = (n, \sum_{i=1}^{n} x_i^\eta) \] ...3.2.5.8

The Weibull distributions do not form an exponential family but belong to a semi-exponential family of dimension 1 in equation (3.2.5.3)
\[ G_1(\theta_1/\theta_2) = \zeta, \ G_2(\theta_2) = \eta \]
\[ H(x/\theta_2) = \sum_{i=1}^{n} x_i^{\eta-1} \]
\[ \psi_1(\theta_1/\theta_2) = \zeta, \ T_1(x/\theta_2) = \sum_{i=1}^{n} x_i^\eta \]

A conjugate family for \( \zeta \) and \( \eta \) is
\[ \pi(\zeta, \eta) = \frac{a^{r}}{\Gamma(r)} \eta^{\eta-1} e^{-a\eta} \frac{\alpha_0^{\alpha_0+1}}{\Gamma(\alpha_0+1)} \zeta^{\alpha_0} e^{-\zeta \alpha_1} \] ...3.2.5.10

and
\[ P(\Theta_2) = p(\eta) = \frac{a^{r}}{\Gamma(r)} \eta^{\eta-1} e^{-a\eta} \] ...3.2.5.11

which is a Gamma distribution with hyperparameters \( \alpha \) and \( r \).

3.2.6 **Semiparametric Inference** :

Often it is impossible to test parametric assumptions of the life time under usual stress \( S_u \).
Therefore it is desirable to make inference on the life time distribution \( F(t/S_u) \) without too specific assumptions.
One approach is to assume certain parametric models only on the accelerating stress levels $S_1, S_2, \ldots$ and to try to obtain estimations for $F(t/S_u)$ in a nonparametric way.

The model is the following. Let $F(t/S)$ denote the c.d.f. of the life time under stress $S$. The relation between the c.d.f. $F(t/S_u)$ of the life time under usual stress $S_u$ and the c.d.f. $F(t/S)$ of the life time under stress $S > S_u$ is given by

$$F(t/S) = F(a(t)/S_u) \text{ for all } t > 0$$

Here $a(.)$ is called the acceleration function (see figure 1).

For linear acceleration functions

$$a(t) = a(s) \cdot t$$

and for power type acceleration functions

$$a(t) = a(s) \cdot t^\beta(s)$$

a nonparametric analysis is developed by Viertil (1987).

In the special case of linear acceleration functions in the considered range of the stress $S$ the so-called relative acceleration coefficient $\alpha_{1,2}$ between two
\begin{align*}
  F(t/S) &= F(\alpha, t/S) \quad \text{for } t > 0 \quad \ldots 3.2.6.4 \\
  \text{For three stress levels } S_1 < S_2 < S_3 \text{ with relative} \\
  \text{acceleration coefficients } \alpha_{ij} \text{ in} \\
  F(t/S_j) &= F(\alpha_{ij} t/S_1) \text{ with } i < j \quad \ldots 3.2.6.5 \\
  \text{we obtain for invertible c.d.f } F(.) \\
  \alpha_{1,3} &= \alpha_{1,2} \alpha_{2,3} \quad \ldots 3.2.6.6 \\
  \text{If the stress level has to be pointed out we write } \\
  \alpha(S_1, S_2) \text{ for } \alpha_{1,2}, \text{i.e.,} \\
  F(t/S_2) &= F(\alpha(S_1, S_2) t/S_1) \quad \text{for } t > 0 \quad \ldots 3.2.6.7 \\
  \text{Now from} \\
  f(t/S + \Delta S) &= F(\alpha(S, S + \Delta S) t/S) \quad \ldots 3.2.6.8 \\
  \text{Using (3.2.6.6) and taking the limit } \Delta S \to 0 \text{ a} \\
  \text{differential equation for } \alpha(s) \text{ is obtained} \\
  \frac{d \alpha(s)}{ds} &= \alpha(s) \lim_{\Delta S \to 0} \frac{\alpha(S, S + \Delta S) - 1}{\Delta S} \quad \ldots 3.2.6.9 \\
  \text{Defining the infinitesimal characteristic} \\
  \delta(s) &= \lim_{\Delta S \to 0} \frac{\alpha(s, s + \Delta s) - 1}{\Delta s} \quad \ldots 3.2.6.10 \\
  \text{The solution of (3.2.6.9) is given by} \\
  \alpha(s) &= \exp \left[ \int_{s_0}^{s} \delta(s) ds \right] \quad \ldots 3.2.6.11 \\
  \text{The infinitesimal characteristic } \delta(s) \text{ is a material} \\
  \text{characteristic depending on the problem under consideration.}
\end{align*}
From $F(t/s_2) = F(\alpha t/s_1)$ it follows for the corresponding densities

$$f(t/s_2) = \alpha f(at/s_1) \quad \text{for } t > 0 \quad \ldots 3.2.6.16$$

and the density $g(u/\alpha)$ of $u_1$ is given by

$$g(u/\alpha) = \int_0^\infty \frac{z}{\alpha} f(\frac{uz}{\alpha}/s_1) f(z/s_1) \, dz \quad \ldots 3.2.6.17$$

using data $D = (t_1, 1, \ldots, t_1, n$ and $t_2, 1, \ldots, t_2, n)$ the transformed data $t_{1,1}/t_{2,1}$ can be used for updating the a-priori distribution $\pi(\alpha)$ of $\tilde{\alpha}$. The likelihood is given by

$$L(\alpha;D) = L(\alpha;u_1, \ldots, u_n) = \prod_{i=1}^{n} g(u_i/\alpha) \quad \ldots 3.2.6.18$$

and the a-posteriori distribution $\pi(\alpha/D)$ of $\tilde{\alpha}$ from Bayes' theorem

$$\pi(\alpha/D) \propto \pi(\alpha) \cdot L(\alpha,D)$$

Under the assumption $\delta(s) = \delta \cdot h(s)$ the quantity $\delta$ is described by a stochastic quantity $\tilde{\delta}$. From equation (3.2.6.13) we obtain

$$\tilde{\delta} = \frac{\log \tilde{\alpha}}{H(s_2) - H(s_1)}$$

Denoting $\alpha(s) = \alpha(s_1,s)$ by $\alpha_s$ and modelling the unknown quantity $\alpha_s$ by a stochastic quantity $\tilde{\alpha_s}$ we obtain

$$\tilde{\alpha_s} = e^\delta[H(s) - H(s_1)] \quad \ldots 3.2.6.19$$

using this equation a probability statement for the c.d.f. $F(./s_1)$ can be given assuming the inverse $F^{-1}(./s)$ of $F(./s)$ exists.
3.2.7 Material Degradation Models:

For certain materials the lifetime of component is determined by a critical value of a physical parameter $p$ and the degradation of $p$ in dependence of time $t$ is given by a differential equation

$$\frac{dp}{dt} = -\frac{C(s)}{\psi(p)} \quad \ldots 3.2.7.1$$

with $C(s) > 0$, $(p) > 0$, where the constant $C(s)$ depends on the applied temperature $S$ (stress). The solution $p(t)$ of equation (3.2.7.1) is given by

$$\int_0^p \psi(p) \, dp = C(s) \cdot (t_0 - t) \quad \ldots 3.2.7.2$$

with

$$t_0 = \frac{1}{C(s)} \int_0^{p(o)} \psi(p) \, dp$$

By the form of equation (3.2.7.1) it follows that the solutions for different values of $p(o)$ are noncrossing shifted decreasing curves as explained in figure 2.

![Figure 2: Solution of eq. (3.2.7.1)](image-url)
In the life time $X$ of the component is determined by a critical value $p_{cr}$ of $p$ and $p_{o,cr}(t)$ is the value of $p(o)$ such that $p(t) = p_{cr}$ we have the following identity

$$[p(t) > p_{cr}] \iff [p(o) > p_{o,cr}(t)] \quad \ldots \ldots 3.2.7.3$$

The function $p_{o,cr}(t)$ is determined by

$$t = \frac{1}{C(s)} \int_{p_{cr}}^{p_{o,cr}(t)} \psi(p) \, dp \quad \ldots \ldots 3.2.7.4$$

Describing the physical parameter $p(o)$ by a stochastic quantity $p$ with c.d.f. $G(.)$ by the following equivalence of events

$$[X < t] \iff [\tilde{p} < p_{o,cr}(t)] \quad \ldots \ldots 3.2.7.5$$

the c.d.f $F(.)|s)$ of the life time at stress level $S$ is given by

$$F(t/s) = G(p_{o,cr}(t)) \text{ for } t > 0 \quad \ldots \ldots 3.2.7.6$$

Assuming the distribution $G(.)$ is parametric with density $g(p/\theta)$ is a Bayesian analysis for the life time distribution using only a sample of $\tilde{p}$ is possible. For a-priori distribution $\pi(\theta)$ of $\theta$ and data $(p_1, \ldots, p_n) = D$ by Bayes' theorem the a-posteriori density $\pi(\theta/D)$ is obtained. In order to obtain the life time distribution the a-posteriori predicative density of $\tilde{p}$ is calculated, i.e.

$$g(p/D) = \int g(p/\theta) \pi(\theta/D) \, d\theta \text{ for } p > 0 \quad \ldots \ldots 3.2.7.7$$
The predicative distribution of life time at stress level $s$ is given by equation (3.2.7.6) using a-posteriori predicative c.d.f. $G(.)/D$.

If the solution $p_{o,cr}(t)$ of equation (3.2.7.4) is of the form $\Psi[C(s),t]$ the stress dependence of the life time distribution is

$$F(t/s) = G(\Psi[C(s),t]/D) \quad \text{for } t > 0 \quad \ldots 3.2.7.8$$

for two accelerating stress levels $s_1 < s_2$ under the assumption that both functions $\Psi(.)$ and $G(.)/D$ are invertible it follows for the acceleration function $a(t)$ in

$$F(t/s_2) = F(a(t)/s_1) \quad \ldots 3.2.7.9$$

that

$$a(t) = \frac{C(s_2)}{C(s_1)} \cdot t \quad \ldots 3.2.7.10$$

Therefore in this case the acceleration function is linear and the Bayesian methods for linear acceleration functions can be applied.

3.2.8 Other Bayesian Approaches in ALT:

There are different proposals for the analysis of accelerated life tests using Bayesian inference. One is to use Kalman Filter Techniques (compare[28]) and another is more pragmatic approach using failure rates[42]. Also the problem of experimental design was considered from a Bayesian viewpoint (compare [9]).
3.2.8. Kalman Filter Approach:

The technique for the time series analysis can be used in ALT in the following way. Replacing the time axis of time series by the (one dimensional) stress $s$, the estimation of life time distribution under design stress $s_u$ is considered as prediction problem. Contrary to time series in ALT the prediction is from large values of $s$ to small values. This is depicted in figure 3.

![Figure 3: Stress pattern from high stress levels to low stress $s_u$.](image)

The recursive process of Kalman Filter for observations $y_t, y_{t-1}, \ldots, y_1$ at time $t, t-1, \ldots, 1$ depending on a parameter $\Theta_t$ is given by the system equation

$$\Theta_t = G_t \Theta_{t-1} + u_t$$  ...3.2.8.1

where $G_t$ is a known quantity and $u_t \sim N(0, u_t)$, with $u_t$ also known, and by the observation equation.

$$y_t = F_t \Theta_t + v_t$$  ...3.2.8.2

with $F_t$ also known. The observation error $v_t$ is assumed to be normally distributed with mean zero and a known variance $v_t$, i.e., $v_t \sim N(0, v_t)$. The dynamic nature
is modelled by the evolution of the parameter $\theta_t$. Denoting $y_t = (y_{t}, \ldots y_{1})$ the inference about $\theta_t$ is carried out by an application of Bayes' theorem:

$$\pi(\theta_t | y_t) \propto \pi(\theta_t/y_{t-1}) \, \mathcal{L}(y_t/\theta_t, y_{t-1}) \ldots 3.2.8.3$$

The expression on the left and the first expression on the right side of the relation are the a-posteriori and the a-priori distribution respectively. The second expression on the right hand side is the likelihood.

Depending on the physical model for the relationship $\theta(s) = \Psi(s, C)$ between a statistical parameter $\theta$ and the stress $s$ the stress dependence in the framework of the Kalman filter is

$$\theta_t = \theta(s_t) = \Psi(s_t, C) \ldots 3.2.8.4$$

From this equation under certain conditions the system equation for the corresponding Kalman Filter can be obtained.

3.2.8.2 A Pragmatic Failure Rate Approach:

For $k$ accelerating stress levels $s_1, \ldots s_{k}$ and usual stress $s_u$ with

$$s_1 > s_2 > \ldots > s_k > s_u$$

and corresponding failure rates $\lambda_u(.)$ and $\lambda_j(.)$, $j=1, \ldots, k$ it is assumed that

$$\lambda_1(t) > \lambda_2(t) > \ldots > \lambda_k(t) > \lambda_u(t) \text{ for } t > 0 \ldots 3.2.8.5$$

Using failure data obtained under stress levels $s_1, s_2, \ldots, s_k$ one wishes to obtain estimates $\hat{\lambda}_j(t)$ for
\( \lambda_j(t), j = 1, \ldots, k, \) such that for same interval \([0,M]\), \( M < \infty \) and all \( t \in [0,M] \) the following stochastic inequalities hold
\[
\hat{\lambda}_1(t) \geq \hat{\lambda}_2(t) \geq \ldots \geq \hat{\lambda}_k(t) \quad \ldots 3.2.8.6
\]

In order to obtain estimates \( \hat{\lambda}_j(t), j = 1, \ldots, k, \) which satisfy (3.2.8.6) is incorporated as prior assumption and using Beta-distributions as a-priori distributions for the probabilities of failure times in certain subintervals of \([0,M]\).

The ultimate interest is to estimate the failure rate \( \lambda_u(\cdot) \) under usual stress \( s_u \). This can be done by weighted averaging of the Bayes estimators at stress levels \( s_j, j = 1, \ldots, k. \)

3.2.8.3 Experimental Design in ALT:

The determination of accelerating stress levels \( s_1, \ldots, s_k \) in order to obtain optimal information concerning the distribution of the life time under usual stress \( s_u \) is an important problem. There are only few papers on optimal design in ALT. Especially in the Bayesian context two papers should be mentioned.

In connection with small samples the paper by Martz and Waterman (1978) which deals with the problem of the determination of the optimal test stress for a single test unit.
Related to so called partially accelerated life tests, i.e., observations are available also under usual stress $s_u$, a paper by DeGroot and Geol (1979) gives some help for constructing optimal designs using cost functions.

*****
CHAPTER - IV

INFERENCE FROM ACCELERATED LIFE TESTS

4.1 A BRIEF SURVEY OF THEORY OF ACCELERATED LIFE TESTS:

4.1.1 Introduction:

Often in reliability studies, accelerated testing is used when one wishes to estimate the life distribution of a device which has a rather long life time under normal use. Items are subjected to greater stress than that of the use condition, and from the observed failure data an estimate of the distribution function of the lifetime of the item under the use condition is formed.

Several parametric and nonparametric procedures have been exposed for estimation from accelerated life test data. In section (4.1.3) and (4.1.4) some of the representative results that have been obtained in the parametric, partially nonparametric, and non-parametric setting will be mentioned briefly.

4.1.2 Notation:

Let $F_V$: The life distribution of items on test under stress level $V$.

$V_1, \ldots, V_k$: $K$ stress levels under accelerated testing, also are in increasing order.

$V_0$: the normal use stress.

$X_1^0$: the true lifetime random variable at stress level $V_1$, $i = 0,1,2,\ldots, k$.

$F_{V_1}$: Probability distribution of $X_1^0$. 
\( x_{ij}^{0}, \ j=1,2,\ldots,n_1 \) : lifetimes of \( n_1 \) random items that are placed on life test at stress level \( V_i \) for each \( i = 0,1,2,\ldots,k. \)

\[ N = \sum_{i=1}^{k} n_i \] : the total number of items.

If the life tests are continued until all \( n_i \) items fail at the stress level \( V_i \), then

\( x_{i1}^{0}, \ldots, x_{in_i}^{0} \) : The failure times for each \( i \).

\( x_{ij}^{(0)}, \ j=1,\ldots,n_1 \) : complete sample of failure times at stress \( V_i \).

In many accelerated life testing situations, some items at one or more of the stress levels may be removed from test at various times due to experimenter's control, or some of the tests at the various levels of stress may be terminated before all items fail.

In these circumstances a censored sample is observed and is called arbitrarily right censored sample. It is assumed that under stress \( V_i \) there is a sequence of constant or random variables \( U_{i1},\ldots,U_{in_i} \) which censors \( x_{i1}^{0},\ldots,x_{in_i}^{0} \) respectively, from the right. That is the observations are the pairs \((x_{ij}, \Delta_{ij})\), \( j=1,2,\ldots,n_i \), where

\[ x_{ij} = \min \{ x_{ij}^{0}, U_{ij} \}, \]

\[ \Delta_{ij} = \begin{cases} 1, & x_{ij}^{0} < U_{ij} \\ 0, & x_{ij}^{0} > U_{ij} \end{cases} \]
Thus $X_{1j}$'s are known as right censored. The type of censorship depends on the nature of the $U_{ij}$'s:

(i) If the $U_{ij}$'s are all constants, then the $X_{1j}$'s are time-censored, if $U_{ij} = C_i$, a constant for all $j = 1, \ldots, n_1$, then the observed sample at stress level $V_i$ is said to be a Type I Censored sample.

(ii) If $U_{ij} = X^o_{i}(r_i)$, the $r_i$th order statistic among $X^o_{ij}, \ldots, X^o_{in_i}$ for all $j = 1, \ldots, n_1$, then the sample is Type II censored sample. That is, the life test at stress level $V_i$ is terminated after $r_i$ failures has occurred.

(iii) If for each $i$, the $U_{ij}$'s are iidrv having common distribution function $H_i$ and independent of $X^o_{11}, \ldots, X^o_{in_i}$, then the observed sample $(X_{ij}, \Delta_{ij}), j = 1, \ldots, n_1$, is a randomly right censored sample. In this case, it is also assume that the censoring random variables $U_{1i}, \ldots, U_{ni}$, $i = 0, 1, \ldots, k$, are independent that is, censoring at one stress level has no effect on the censoring at other stress levels.

In case of right-censorship, the random variable $X^o_{ij}$ and $U_{ij}$ for fixed $i$ and $j$ are called life times of the $j^{th}$ item due to two independent causes of failure or independent competing risks. The lifetime variables $X^o_{ij}$ denotes the lifetime due to the 'main' cause of failure, and $U_{ij}$ denotes
the lifetime due to the other independent competing
cause. For more than two independent competing risk see
Basu and Ebrahimi (1982).

4.1.3 Some Parametric Procedures:

In this section we will be illustrated with the simplest
parametric case of an exponential life distribution at each
stress level and the power rule model as the acceleration
model. Also, some results of Shaked (1978) will be stated
that generalize the power rule model and Arrhenius model
with life distributions which have hazard rate functions
of the form \( r(t) = Ag(t) + Bh(t) \). This class of life
distributions contains the exponential distributions as
well as other distributions which have been considered in
the literature. Results for Weibull and other life
distributions can be found, for example, in papers by
Mann (1972), Nelson (1972,1975), and Nelson and Meeker
(1978), and will not be discussed here.

Suppose that life tests are to be conducted at \( K \)
accelerated stress levels, \( V_1, i = 1,2,\ldots,k \), on devices,
which are assumed to have an exponential life distribution
with parameter \( \lambda_i = 1/\mu_i \), where \( \mu_i \) is the mean lifetime
under stress \( V_i \). In order to assume an independence among
K life tests, a randomization procedure is performed as
follow: choose a value \( V_j \) at random from the \( K \) values
\( V_i, i = 1,\ldots,k \), and put \( n_j \) items on life test, which is
conducted under constant application of stress $V_j$. The test is termination after $r_j$ failure have occurred and respective times to failure, $t_{1j}, t_{2j}, \ldots, t_{r_jj}$ are recorded. Next choose another stress $V_1$ at random from remaining $K-1$ stress levels and repeat the procedure. Continue this procedure until all $k$ life tests have been performed, yielding as data, the set $[V_1, n_1, r_1, \hat{\mu}_1], i = 1, \ldots, k$, where $\hat{\mu}_1$ is an estimator of $\mu_1$.

The unique minimum variance unbiased estimator of $\mu_1$ is (Epstein and Sobel, 1954)

$$\hat{\mu}_1 = \frac{r_1}{\sum_{j=1}^{r_1} t_{j1} + (n_1 - r_1) t_{r_1j}}/r_1 \quad \ldots 4.1.3.1$$

The p.d.f. of $\hat{\mu}_1$ is a gamma density function with a shape parameter $\lambda_1$ scale parameter $r_1/\mu_1$, so that

$$g(\hat{\mu}_1) = \frac{(r_1/\mu_1)^{\lambda_1}}{(\lambda_1)^{r_1}} e^{-r_1\hat{\mu}_1/\mu_1} (\hat{\mu}_1)^{r_1-1}, \quad 0 < \mu_1 < \infty \quad \quad r_1 > 1 \quad \quad \ldots 4.1.3.2$$

Consider power rule model

$$\mu_1 = \alpha/V_1^\beta, \quad \alpha > 0 \text{ for all } V_1 \text{ within a specified range.}$$

Suppose an accelerated life test is conducted according to the randomization procedure, $\alpha$ and $\beta$ are to be estimated so that inferences can be made about the mean lifetime $\mu_0$ at the normal use stress $V_0$. In order to obtain asymptotically independent estimators of $\alpha$ and $\beta$, Singpurwalla (1971) amended
the power rule model slightly, without changing its basic character, as

$$\mu_i = \alpha \left( \frac{V_i}{\hat{V}} \right)^{\beta}$$

for all value of $i$, where $\hat{V}$ is the weighted geometric mean of the $V_i$'s. That is,

$$\hat{V} = \frac{1}{k} \sum_{i=1}^{k} V_i^R_i$$

where

$$R_i = \frac{r_i}{k} \sum_{j=1}^{k} r_j$$

The p.d.f of this power rule model is also followed the gamma distribution with shape parameter $r_1$ and scale parameter.

Since

$$1/\mu_i = 1/\alpha \left( \frac{V_i}{\hat{V}} \right)^{\beta}$$

$$\Rightarrow r_1/\mu_i = r_1/\alpha \left( \frac{V_i}{\hat{V}} \right)^{\beta}$$

i.e. the scale parameter is $r_1/\alpha \left( V_i/\hat{V} \right)^{\beta}$

Thus p.d.f. $f(\alpha, \beta, \hat{\mu})$ is

$$f(\alpha, \beta, \hat{\mu}) = \frac{[r_1/\alpha \left( V_i/\hat{V} \right)^{\beta}]^{r_1}}{\Gamma(r_1)} e^{-r_1\hat{\mu}/\alpha(V_i/\hat{V})^{\beta}} (\hat{\mu})^{r_1-1}$$

To find maximum likelihood estimators of $\alpha$ and $\beta$, say, $\hat{\alpha}$ and $\hat{\beta}$:

The likelihood function of $\alpha$ and $\beta$ can be written as

$$L(\alpha, \beta, \hat{\mu}) = \frac{[r_1/\alpha \left( V_i/\hat{V} \right)^{\beta}]^{k r_1}}{\Gamma(r_1)^k} \cdot e^{-\sum r_1 \hat{\mu}/\alpha(V_i/\hat{V})^{\beta}} (\hat{\mu})^{k r_1-1}$$

Therefore
\[ \log L(a, \beta; \hat{\mu}) = \log C + \sum_{i=1}^{k} r_i \log \frac{r_i}{a(V_1/\dot{V})^\beta} - \sum_{i=1}^{k} \mu_i a(V_1/\dot{V})^\beta \]

where \( C \) is constant and independent of \( a \) and \( \beta \).

Hence,
\[ \log L(a, \beta; \hat{\mu}) = \log C + \sum_{i=1}^{k} r_i \log \frac{r_i}{a} + \beta \sum_{i=1}^{k} r_i \log (V_1/\dot{V}) - \frac{1}{\alpha} \sum_{i=1}^{k} \mu_i (V_1/\dot{V})^\beta \]

Differentiating with respect to \( a \) and \( \beta \) and equating to zero we get,
\[ \hat{a} = \frac{\sum_{i=1}^{k} r_i \hat{\mu}_i (V_1/\dot{V})^\beta}{\sum_{i=1}^{k} r_i} \quad \ldots 4.1.3.3 \]

and
\[ \sum_{i=1}^{k} r_i \log (V_1/\dot{V}) = \frac{1}{\alpha} \sum_{i=1}^{k} \mu_i (V_1/\dot{V})^\beta \log (V_1/\dot{V}) \quad \ldots 4.1.3.4 \]

since \( \sum_{i=1}^{k} r_i \log (V_1/\dot{V}) = 0 \), it is first necessary to find those values of \( \beta \), say \( \hat{\beta} \) for which
\[ \sum_{i=1}^{k} r_i \hat{\mu}_i (V_1/\dot{V})^\hat{\beta} \log (V_1/\dot{V}) = 0 \quad \ldots 4.1.3.5 \]

Since equation (4.1.3.5) is nonlinear, the solution \( \hat{\beta} \) must be found numerically. Singpurwalla (1971) reported that, using the Newton Raphson method, for various sets of data generated by computer, the solutions converged in five to ten iterations. Then the numerical solution \( \hat{\beta} \) gives a unique value \( \hat{a} \) from (4.1.3.3). Thus, \( \mu_0 \) is estimated by substituting \( \hat{a} \) and \( \hat{\beta} \) into the amended power rule model at the normal use stress \( V_0 \).
It is easy to verify that since

$$-E \left[ \frac{\partial^2 \log L}{\partial \beta^2} \right] = k r_1 \left[ \log \left( \frac{V_1}{\hat{V}} \right) \right]^2,$$

$$-E \left[ \frac{\partial^2 \log L}{\partial \sigma^2} \right] = k \frac{r_1}{\sigma^2},$$

and

$$-E \left[ \frac{\partial^2 \log L}{\partial \beta \partial \sigma} \right] = 0,$$

The asymptotic variances and the asymptotic covariance of \( \hat{\beta} \) and \( \hat{\sigma} \) are

$$\text{Var}(\hat{\beta}) = \sigma^2_{\beta} = \left( \frac{r_1}{\sum \log \left( \frac{V_i}{\hat{V}} \right) \right)^2, \text{ and}$$

$$\text{Var}(\hat{\sigma}) = \sigma^2_{\sigma} = \sigma^2 \left( \frac{r_1}{\sum \log \left( \frac{V_i}{\hat{V}} \right) \right)^2, \text{ and}$$

$$\text{Cov}(\hat{\beta}, \hat{\sigma}) = 0.$$  \(...4.1.3.7\)

Since the maximum likelihood estimators are asymptotically unbiased and distributed as bivariate normal density, \( \hat{\alpha} \) and \( \hat{\beta} \) are asymptotically independent. This was the reason for using the amended failure rule model. These results can be used for asymptotic inferences from the accelerated test data.

Shaked (1978) has studied the accelerated life testing problem for a parametric class of life distribution, with linear hazard rate functions.

Suppose that a device has a life distribution of the form

$$F^0(t) = 1 - \exp \left[ -\alpha G(t) - \beta H(t) \right], \ t \geq 0 \ ...4.1.3.8$$

where \( \alpha \) and \( \beta \) are unknown parameters and \( G \) and \( H \) are known differentiable functions of \( t \) which do not depend on
\(\alpha\) and \(\beta\). Some well known life distributions that have been discussed in the literature are of the form \((4.1.3.8)\), e.g. Brain, 1974, and Prairie and Ostle, 1961.

Assume that the underlying distribution is exponential at each stress level and to let the hazard rate be a function of \(V > 0\). According to the power rule model the hazard rate \(r_1\) under stress \(V_1\) is

\[
r_1 = \alpha V_1^\beta, \quad \alpha > 0, \quad -\infty < \beta < \infty \quad \ldots \quad (4.1.3.9)
\]

and according to the Arrhenius model

\[
r_1 = \alpha \exp \left[-\frac{\beta}{V_1}\right], \quad \alpha > 0, \quad -\infty < \beta < \infty \quad \ldots \quad (4.1.3.10)
\]

Since usually, \(V_1 > V_0\) for \(i \neq 0\). Acceleration means that \(r_1 > r_0\) when \(V_1 > V_0\), where \(r_0\) is the hazard rate under normal use conditions. Shaked (1978) also introduced and acceleration model which generalizes \((4.1.3.9)\) and \((4.1.3.10)\) applied it to the life distribution \((4.1.3.8)\), and obtained the maximum likelihood estimators of the unknown parameters. More explicitly, if \(\omega_1 = q(V_1)\), where \(q\) is a known positive increasing function of \(V > 0\) and assumes that under stress level \(V_1\) the life distribution is

\[
F_{V_1}(t) = 1 - \exp \left[-\omega_1 \left[\alpha G(t) + \beta H(t)\right]\right], \quad t > 0 \quad \ldots \quad (4.1.3.11)
\]

or equivalently that the hazard rate function is

\[
r_1(t) = \omega_1 \left[\alpha g(t) + \beta h(t)\right], \quad t > 0 \quad \ldots \quad (4.1.3.12)
\]

where \(g = G'\) and \(h = H'\).

Note that if \(q(V) = V, G(t) = t\) and \(\beta = 0\) then \((4.1.3.12)\) reduces to \((4.1.3.10)\).
Shaked (1978) showed how to obtain the maximum likelihood estimators, $\hat{a}$, $\hat{\beta}$ and $\hat{p}$ and he expressed the estimators $\hat{a}$ and $\hat{\beta}$ as explicit functions of $\hat{p}$ [$\hat{a} = \hat{a}(\hat{p})$, $\hat{\beta} = \hat{\beta}(\hat{p})$, say]. He found the exact region of $\hat{p}$'s in which $\hat{p}$ must lie. By substituting $\hat{a}(\hat{p})$ and $\hat{\beta}(\hat{p})$ in the likelihood equation, he got a function of one variable $(\hat{p})$ which was to be maximized over the bounded interval to find $\hat{p}$ and with help of $\hat{p}$ he computed $\hat{a}$ and $\hat{\beta}$. The advantage of his procedure was that no more than one equation with one unknown variable must be solved at a time. Then to get the maximum likelihood estimator of the lifetime distribution function under the normal use stress, $V_0$, $\hat{a}$, $\hat{\beta}$ and $\hat{p}$ can be substituted into (4.1.3.11) with $i = 0$, expressions for the asymptotic variance-covariance matrix of $\hat{a}$, $\hat{\beta}$ and $\hat{p}$ involve integrals which must be computed numerically.

4.1.4 Some Nonparametric Methods:

In some situations, it may not be possible to determine a reasonable parametric form for the lifetime distribution of items under test at either the normal use stress or the accelerated stresses. That is, no specifications of a family of parametric life distributions is given. For such situations, inference procedures have been developed which are non-parametric. If the particular form of the acceleration model is not assumed, it is necessary to have at least a few observations from the normal use stress level.
In this section some of the nonparametric methods will be described briefly. These will include both types of methods, those without a specific acceleration model and those with utilize a particular form for acceleration model.

Barlow and Schefer (1971) assumed that data were available from both the normal use life distribution $F_0$ and one accelerated life distribution $F_1$, where the exact forms of $F_0$ and $F_1$ were unspecified. It was assumed that $F_1$ and $F_0$ were related by a time transformation $a(t)$, where

$$F_1(t) = F_0[a(t)]$$

So that, assuming $F_0^{-1}$ exists,

$$a(t) = F_0^{-1}(t)$$

where $a(t)$ was unknown and $a(t) > t$. Also he assumed that $F_1$ and $F_0$ were increasing failure rate average (IFRA) distributions, i.e. if the failure rate $r(t)$ exists, then $\frac{1}{t} \int_0^t r(u)du$ were nondecreasing in $t$.

More generally, a distribution $F$ was IFRA if $(-\log[1-F(t)]) /t$ was nondecreasing in $t > 0$.

The problem considered was as follows:

Let $Y_1 \leq \ldots \leq Y_m$ be ordered observations from $F_0$ (the unaccelerated life distribution) and $X_1 \leq \ldots \leq X_n$ be ordered observations from $F_1$ (the accelerated life distribution). He obtained the empirical distribution
Then obtain estimates $\hat{F}_{om}$ and $\hat{F}_{ln}$ such that

(i) $\hat{F}_{om}$ and $\hat{F}_{ln}$ are IFRA:

(ii) $\hat{F}_{om}(x) < \hat{F}_{ln}(x)$ for all $x$

(iii) $\hat{F}_{om}$ and $\hat{F}_{ln}$ are closest to $F_{om}$ and $F_{ln}$ respectively in a least squares sense.

The estimate $\hat{F}_{om}$ of $F_0$ is obtained using data from both the accelerated and unaccelerated stress levels.

To obtained an estimate of the unknown failure distribution $F_1$, assuming that $-\log \bar{F}_1(t)/t$ is nondecreasing on $[0, \infty]$ where $\bar{F}_1(t) = 1 - F_1(t)$. Suppose $K$ (1 $\leq k $ $\leq n$) failures were observed. Let $X_1 \leq X_2 \leq \ldots \leq X_k$ denote the ordered ages at failure and let $n_k$ denote the number of items surviving just before time $X_k$. The product limit estimate of $F_1$ can be written as

$$
\bar{F}_{ln}(t) = \begin{cases}
1, & 0 < t < X_1 \\
\frac{1}{\pi} \sum_{j=1}^{n_k-1} \left( \frac{n_k-1}{n_k} \right), & X_1 < t < X_{k+1}, 
\end{cases} \quad \text{for } i = 1, \ldots, q-1
$$

where

$$
q = \begin{cases}
k & \text{if } k < n \\
n & \text{if } k = n \quad (\text{no censoring})
\end{cases}
$$

Let $\lambda(t) = -[\log \bar{F}_1(t)]/t$ and obtain an estimate of $\lambda$ by a nondecreasing function $\hat{\lambda}_n$. Let
\[ \lambda_n(X_i) = \frac{-\log F_{ln}(X_i)}{X_i} \quad \text{for } i = 1, 2, \ldots, q \]

and define
\[ \hat{\lambda}_n(X_i) = \max_{s \leq i} \min_{t \leq \gamma} \frac{\sum_{j=s}^{\gamma} \lambda_n(X_j) F_{ln}(X_j) - \sum_{j=s}^{\gamma} F_{ln}(X_j)}{t - \sum_{j=s}^{\gamma} F_{ln}(X_j)} \quad \ldots \text{4.1.4.3} \]

where \( F_{ln}(X_j) = F_{ln}(X_j^-) - F_{ln}(X_j^+) \). Then define
\[ \hat{\lambda}_n(t) = \begin{cases} 0 & , t < X_1 \\ \hat{\lambda}_n(X_1) & , X_1 < t < X_{i+1} , i = 1, \ldots, q-1 \\ +\infty & , t \geq X_q \quad \ldots \text{4.1.4.4} \end{cases} \]

So the IFRA estimate for the distribution function is given by
\[ \hat{F}_{ln}(t) = \begin{cases} 1 & , t < X_1 \\ \exp[-\hat{\lambda}_n(X_1)t] & , X_1 < t < X_{i+1} , i = 1, \ldots, q-1 \\ 0 & , t \geq X_q \quad \ldots \text{4.1.4.5} \end{cases} \]

Now, to estimate \( F_0 \) and \( F_1 \), Suppose \( X_1 < \ldots < X_k (\leq n) \) failures from a sample of size \( n \) for a distribution \( F_1 \), and \( Y_1 < Y_2 < \ldots < Y_r (\leq m) \) from a sample of size \( m \) for a distribution \( F_0 \). Assume that

(i) \( F_1 \) and \( F_0 \) are IFRA

(ii) \( F_1(t) > F_0(t) \) for all \( t > 0 \)

Let
\[ \lambda_n(X_i) = -\log \frac{F_{ln}(X_i)}{X_i} \quad \text{and} \quad \gamma_m(Y_i) = -\log \frac{F_{om}(Y_i)}{Y_i} \quad \text{and} \]

...
\( \hat{\lambda}_1, \hat{\gamma}_1 \) denote estimates of \( \lambda(x_1) \) and \( \gamma(y_1) \), respectively. To determine the \( \hat{\lambda}_1 \) and \( \hat{\gamma}_1 \), so that,

\[
\begin{align*}
\min & \quad \sum_{i=1}^{k} [\hat{\lambda}_1 - \hat{\lambda}_n(x_i)]^2 F_{\text{ln}}(x_i) + \sum_{j=1}^{r} [\hat{\gamma}_1 - \hat{\gamma}_m(y_j)]^2 F_{\text{om}}(y_j) \\
\text{subject to} & \quad 0 \leq 1 \leq 2 \leq \ldots \leq k \\
& \quad 0 \leq \gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_r \\
& \text{and} \quad \hat{\lambda}_{\text{om}}(t) \geq \hat{\lambda}_{\text{ln}}(t)
\end{align*}
\]

This problem can be solved by quadratic programming, and the solution is the best under the constraints in a least squares sense.

Shaked, Zimmer, and Bell (1979) assumed a general relationship between the life distribution of any two different stress levels as an acceleration model. Specifically, for every pair of stress levels \( V_i, V_j \), \( i \neq j, i, j = 0, 1, \ldots, k \), it was assumed that a known function \( m \) existed so that the life distributions satisfied

\[
F_{V_j}(t) = F_{V_1}(m(a, V_j, V_1, t)), \quad t > 0 \quad \ldots 4.1.4.7
\]

where \( F_V \) denoted the d.f. of the lifetime of a device subject to stress \( V \) and \( a \) was an unknown parameter. Let \( x_{1i}^0, i = 1, \ldots, n_i, i = 1, \ldots, k \) be observations under stress levels \( V_i, i = 1, \ldots, k \). Hence, no observations are required for stress \( V_0 \). Here \( K, V_i \)'s, and \( n_i \)'s are
selected and fixed in advance, with no censoring of the data.

For simplicity, he took

\[ m(\alpha, V_j, V_1, t) = \left[ g(\alpha, V_j)/g(\alpha, V_1) \right] t \quad \ldots.4.1.4.8 \]

where \( g \) was a known function which was positive at each \( V \) and nondecreasing in \( V \).

Their procedure was

(i) to estimate \( \hat{\alpha} \) and then

(ii) for every \( t > 0 \), to estimate \( F_{V_0}(t) \) from the data and from the estimate \( \hat{\alpha} \).

(iii) The estimate of \( F_{V_0} \) was then used to estimate the means life time, the percentiles, and the median in the normal use environment.

An estimate of \( \alpha \) was obtained by first estimating the scale factor between \( F_{V_j} \) and \( F_{V_1} \), the factor was

\[ \Theta_{ij} = g(\alpha, V_j)/g(\alpha, V_1) \quad , \quad i \neq j \quad , \quad \ldots.4.1.4.9 \]

Any consistent estimator of \( \Theta_{ij} \) could be used, such as

\[ \hat{\Theta}_{ij} = \frac{\bar{X}_1^o}{\bar{X}_j^o} \], where \( \bar{X}_1^o = 1/\eta_1 \sum_{i=1}^{n_1} X_{1i}^o \)

Then \( \alpha \) was estimated by a weighted average

\[ \hat{\alpha} = \sum \omega_{ij} W_{ij} \hat{\Theta}_{ij} \]

where \( \omega_{ij} \)'s were chosen depending on \( g \) and \( \hat{\Theta}_{ij} \), and \( \hat{\Theta}_{ij} \) was obtained from solving \( \hat{\Theta}_{ij} = g(\hat{\alpha}_{ij}, V_j)/g(\hat{\alpha}_{ij}, V_1) \).
Using $\hat{\alpha}$, the observations from the K accelerated tests were rescaled by

$$\tilde{X}_{1i} = \left[ \frac{g(\hat{\alpha}, V_1)}{g(\hat{\alpha}, V_0)} \right] X_{1i}, \quad i = 1, \ldots, n_1, \quad i = 1, \ldots, k$$

Finally, the empirical distribution function $\hat{F}_{V_0}$ of these $N = \sum n_i$ rescaled observations was used as an estimate of $F_{V_0}$.

Basu and Ebrahimi (1982) extended the results of Shaked, Zimmer, and Ball (1979) and of Shaked and Singpurwalla (1982) to the case of randomly right-unscored samples $(X_{1i}, A_{1i})$ at each accelerated stress $V_1$. The choice of the acceleration model $m$ was a power rule model. That is, $m(\alpha, V_j, V_1, t) = \left[ \frac{g(\alpha, V_j)}{g(\alpha, V_1)} \right] t = \left( \frac{V_j}{V_1} \right)^{\alpha} t$ was used. The estimator of the scale factor $\hat{\theta}_{1j} = \left( \frac{V_j}{V_1} \right)^{\alpha}$ was that obtained for the two sample problem from arbitrarily right-unscored samples by Padgett and Wei (1982). The value of $\alpha$ was estimated in a manner suggested by Shaked and Singpurwalla (1982), and the estimate was used to rescale the censored observations from the accelerated tests by $\tilde{X}_{1i} = \left( \frac{V_1}{V_0} \right)^{\alpha} X_{1i}, \quad i = 1, \ldots, n_1, \quad i = 1, \ldots, k$, and obtained the estimator for $F_{V_0}$. Basu and Ebrahimi (1982) also extended the results of Shaked and Singpurwalla (1982) to the case of I-component series systems at each accelerated stress level.
Shaked and Singpurwalla (1982) combined the techniques of Sethuraman and Singpurwalla (1982) and of Shaked, Zimm, and Ball (1979) for complete samples from each accelerated stress. The rescaling technique using the estimated \( \alpha \) from the power rule model for the scale change parameters described above was used to estimate \( F_{Y_0} \).

Shaked and Singpurwalla (1983) considered accelerated life testing wherein the stress on an unfailed item can increase at a preassigned test time. Such tests are called step-stress tests. Their approach was nonparametric in the sense that no assumption were made about the underlying lifetime distribution at each stress. Their model for step-stress testing was based on the ideas of shock models and wear processes.

4.2 SOME ASPECTS OF ACCELERATED LIFE TESTING BY PROGRESSIVE STRESS:

4.2.1 Introduction:

Accelerated life testing over stresses specimens to make them fail sooner. There are three kinds of stress: constant, step, and progressive.

1) Constant stress:

Specimen are tested at several constant stresses (higher than design stress). Life at a design stress is estimated by regression method (Mann, Schafer, Singpurwalla, 1974).
ii) **Step Stress**: Stress on the specimens is a planned step function. The stress starts at a low level and is held a specified time, then the stress level is 'increased and held' repeatedly until the specimen fails. One uses the test data and the model, which relates the c.d.f. of product life under constant stress to the c.d.f. under step stress, to estimate product reliability.

iii) **Progressive Stress**: Progressive stress is similar to step stress, but the stress on specimens is a progressive (continuous) function.

Here, the aim is to see the relationship between the cdf under constant stress and the cdf under progressive stress. This paper, based on Nelson (1980) and others. It starts with some basic assumptions and then derives the cdf under progressive stress, when the cdf under design constant stress is an exponential or a weibull distribution.

The paper discusses the special case when the acceleration equation satisfies inverse power law and the progressive stress is directly proportional to time. For this some of the maximum likelihood estimator properties are discussed. This is based on the work of Nelson (1980) and others.
4.2.2 Notation:

- $\alpha$: rate parameter of Weibull distribution
- $S, S_1, S(t), S(t)$: Stress
- $t_i$: time of sample $i$, or, time at which the stress goes from $S_i$ to $S_{i+1}$.
- $F(t, S), \tilde{F}(t), F(t)$: CDF of product life under constant, step, progressive stress.
- $C_i = \max (t_i - t_{i-1})$
- $a, b$: parameters of function that relates rate parameter and stress
- $\beta$: Shape parameter of Weibull distributions implies an estimate.
- $K$: slope of linear progressive stress.
- $N$: Sample size.

4.2.3 Basic Assumptions:

i) For design constant stress, the life distribution is a Weibull distribution or an exponential distribution.

ii) If the life has a Weibull distribution or an exponential distribution under design constant stress, then under other constant stresses, the distribution is the same type.

iii) The failure mechanisms of a product are the same at all levels of stress.

iv) The life and stress have a known relationship.

v) If held at the current stress, survivors will fail accordingly to the CDF for that stress but starting at the previously accumulated fraction failed.
vi) The testing is complete sample testing and the samples are arranged from small to large.

vii) The acceleration equation satisfies an inverse power law.

viii) Progressive stress is directly proportional to time.

The model was put forward by Nelson (1980) and others, which relates the cdf under step stress to the cdf under constant stress.

We assume that, for the Weibull distribution, different cdfs are determined by different rate parameters, viz., the rate parameter depends only on the current stress and does not remember stress history. The shape parameter is assumed to be a constant.

That is, the model means that if the cdf under constant stresses $S_1 < S_2 < \ldots < S_k$ are $F(t, S_1)$, $F(t, S_2)$ $\ldots F(t, S_k)$, then under step stress.

$$ S(t) = S_i, \quad t_{i-1} \leq t < t_i, \quad t_0 = 0, \quad i = 1, 2, \ldots, k $$

the cdf is

$$ F(t) = F \left[ (t-t_{i-1}) + T_{i-1}, S_i \right], \quad t_{i-1} \leq t < t_i, \quad t_0 = 0 $$

$$ i = 1, 2, \ldots, k $$ .4.2.3.1

where $T_{i-1}$ is determined by

$$ F(T_{i-1}, S_i) = F(t_{i-1}-t_{i-2} + T_{i-2}, S_{i-1}) \quad i = 1, 2, \ldots, k, \quad T_0 = 0 $$

$$ \ldots .4.2.3.2 $$
4.2.4 The cdf under accelerated life test in by progressive stress:

Suppose that the product is under accelerated life testing by progressive stress $S(t)$. By assumption (i), the cdf under design constant stress $S$ is:

$$F(t, S) = 1 - e^{-(at)^\beta}, \quad t > 0$$

...4.2.4.1

For any progressive stress $S(t)$, there is a step stress

$$\tilde{S}(t) = S(t_i); \quad t_{i-1} < t < t_i, \quad t_0 = 0, \quad i = 1, 2, \ldots, k$$

...4.2.4.2

where $t_i$ are points in the domain of $S(t)$. $\tilde{S}(t)$ is an approximation of $S(t)$, and

$$S(t) = \lim_{\delta \to 0} \tilde{S}(t)$$

...4.2.4.3

Assumption (ii) gives the cdf's under constant stresses

$$S(t_1) < S(t_2) < \ldots < S(t_k)$$

as

$$F(t, S(t_i)) = 1 - e^{-(a_i t)^\beta_i}, \quad t_{i-1} < t < t_i, \quad t_0 = 0, \quad i = 1, \ldots, k$$

...4.2.4.4

and by assumption (iii) we assume that the shape parameter of cdf is invariant under different stresses.

i.e.

$$\beta = \beta_1 = \beta_2 = \ldots = \beta_k$$

Then (4.2.4.4) becomes

$$F(t, S(t_i)) = 1 - e^{-(a_i t)^\beta}, \quad t_{i-1} < t < t_i, t_0 = 0$$

...4.2.4.5

From assumption (V), the cdf under step stress $\tilde{S}(t)$ is

$$F(t) = F[(t-t_{i-1})+T_{i-1}, S(t_i)]; \quad t_{i-1} < t < t_i, \quad t_0 = 0, \quad i = 1, 2, \ldots, k$$

...4.2.4.6
where $T_{1-1}$ is an equivalent age and is determined by

$$F[T_{1-1}, S(t_1)] = F[t_{1-1} - t_{1-2} + T_{1-2}, S(t_{1-1})]$$

$T_0 = 0, \ i = 1, 2, \ldots, k \ldots 4.2.4.7$

Combine (4.2.4.5) and (4.2.4.7), we have

$$F[T_{1-1}, S(t_1)] = 1 - e^{-\left(\alpha_1 T_{1-1}\right)^{\beta}}$$

and

$$F[t_{1-1} - t_{1-2}, S(t_{1-1})] = 1 - e^{-\left(\alpha_{1-1}(t_{1-1} - t_{1-2} + T_{1-2})\right)^{\beta}}$$

From (4.2.4.7) we have

$$1 - e^{-\left(\alpha_1 T_{1-1}\right)^{\beta}} = 1 - e^{-\left(\alpha_{1-1}(t_{1-1} - t_{1-2} + T_{1-2})\right)^{\beta}}$$

or,

$$\alpha_1 T_{1-1} = \alpha_{1-1}(t_{1-1} - t_{1-2} + T_{1-2}) \quad \ldots 4.2.4.8$$

$$\Rightarrow T_{1-1} = \left[\alpha_{1-1}(t_{1-1} - t_{1-2})\right]/\alpha_1 + \left[\alpha_{1-1} T_{1-2}\right]/\alpha_1$$

Now from (4.2.4.8) we have

$$T_{1-1} = \left[\alpha_{1-1}(t_{1-1} - t_{1-2})\right]/\alpha_1 + \left[\alpha_{1-2}(t_{1-2} - t_{1-3} + T_{1-3})\right]/\alpha_1$$

Finally we have

$$T_{1-1} = \sum_{n=1}^{1-l} \frac{\alpha_{1-n}(t_{1-n} - t_{1-n-1})}{\alpha_1}$$

$$= \sum_{n=1}^{1-l} \frac{\alpha_n(t_n - t_{n-1})}{\alpha_1} \quad \ldots 4.2.4.9$$

Assumption (IV) gives

$$\alpha = \Phi(S)$$

where $S$ is a function of time. Therefore $\alpha$ is also a function of time.
\[ a = \Phi[S(t)] = \alpha(t) \] ...4.2.4.10a
\[ a_1 = \Phi[S(t_1)] = \alpha(t_1) \] ...4.2.4.10b

Now, combine (4.2.4.9) into (4.2.4.6) we have

\[ \hat{F}(t) = 1 - e^{-[\alpha_1(t-t_{1-1}) + T_{1-1}]}^\beta \]
\[ = 1 - e^{-[\alpha_1(t-t_{1-1}) + \sum_{n=1}^{1-1} \alpha_n(t_n-t_{n-1})]}^\beta \]

From (4.2.4.10b), \( \hat{F}(t) \) becomes

\[ \hat{F}(t) = 1 - e^{-[\alpha(t_1)(t-t_{1-1}) + \sum_{n=1}^{1-1} \alpha_n(t_n-t_{n-1})]}^\beta ...4.2.4.11 \]

\[ t_{1-1} < t < t_1, \ t_0 = 0, \ i = 1, 2, ... \]

If \( \xi \to 0 \), then \( i \to \infty \) and \( t_{1-1} \to t \). If \( \xi \to 0 \), \( i \to \infty \), then

\[ (t_n - t_{n-1}) \to 0, \ \text{and} \ \sum_{n=1}^{1-1} \alpha(t_n)(t_n-t_{n-1}) \to \int_0^t \alpha(t') dt', \]

that is the sum becomes integral. Since the step stress \( S(t) \) becomes progressive stress \( S(t) \) when \( \xi \to 0 \), therefore the cdf under step stress converges to the cdf under progressive stress when \( \xi \to 0 \). That is

\[ F(t) = \lim_{\xi \to 0} \hat{F}(t) = 1 - e^{-[\int_0^t \alpha(t') dt']}^\beta ...4.2.4.12 \]

**Theorem**: If the distribution under constant stress \( S \) is Weibull

\[ F(t, S) = 1 - e^{-\alpha(t)^\beta}, \ t > 0 \]

then, the cdf under progressive stress \( S(t) \) is

\[ F(t) = 1 - e^{-\left[ \int_0^t \alpha(t') dt' \right]^\beta} \]

where \( \alpha(t) = \Phi[S(t)] \), and \( \alpha = \Phi(S) \)
The cdf under progressive stress can be obtained by replacing \( a t \) in the cdf under constant stress with \( t \int_a(t')dt' \). Constant stress and step stress are particular cases of progressive stress. Equation (4.2.4.12) is valid for constant stress and step stress, too.

If \( S(t) \) becomes constant stress \( S \) then
\[
F(t) = 1 - e^{-\int_0^t \alpha dt'} = 1 - e^{-\alpha t}^\beta
\]
which is same as eqn. (4.2.4.1)

If the stress \( S(t) \) becomes a step stress \( S(t) \), then
\[
F(t) = 1 - e^{-\int_0^t \Phi(S(t'))dt'} = 1 - e^{-\alpha(t_1)(t-t_{i-1}) + \sum_{n=1}^{i=1} \alpha(t_n)(t_n-t_{n-1})}^\beta
\]
\[t_{i-1} < t < t_i, \quad t_0 = 0, \quad i = 1, 2, \ldots\]
which is the same form as \( F(t) \) is (4.2.4.11)

**Corollary**: Let the distribution under design constant stress be Weibull, and the rate parameter and stress be related as
\[
a = a_S^b \quad \text{(by assumption (vii))}
\]
and stress can be
\[
S(t) = kt \quad \text{(by assumption (viii))}
\]
where \( a > 0, \quad b > 0, \quad k > 0 \). Then the cdf under progressive stress \( S(t) \) is given as
This can be written as

\[ F(t) = 1 - e^{-\left[\frac{ak^{b}t^{b+1}}{(b+1)}\right]^\beta}, \quad t > 0 \quad \ldots 4.2.4.13 \]

Therefore, the cdf under linear progressive stress is a Weibull distribution with new scale and shape parameter, that is

\[ a^* = \frac{ak^b}{(b+1)} \]
\[ \beta^* = (b+1)\beta \]

The exponential distribution is a special case of Weibull distribution when \( \beta = 1 \).

**4.2.5 Estimates of Parameters**:

1) Estimating parameters \( a, b, \beta \) of the distribution:

By the assumption (VI), with cdf (4.2.4.12). We have the pdf, the likelihood function and its derivations with respect to \( a, b \) and \( \beta \).

That is

\[ f(t) = \beta(ak^b)^\beta t^{(b+1)\beta-1} e^{-\left[\frac{ak^{b}t^{b+1}}{(b+1)}\right]^\beta/(b+1)^{\beta-1}} \quad \ldots 4.2.5.1 \]

\[ L(a, b, \beta) = N \prod_{i=1}^{N} \beta(ak^b)^\beta t_i^{(b+1)\beta-1} e^{-\left[\frac{ak^{b}t_i^{b+1}}{(b+1)}\right]^\beta/(b+1)^{\beta-1}} \quad \ldots 4.2.5.2 \]
\[ \log L = N \log \beta + N \beta \log a + N \beta \log K + \left[ (b + 1) \beta - 1 \right] \]

\[ \sum_{i=1}^{N} \log t_i - N (\beta - 1) \log (b+1) - \sum_{i=1}^{N} \left[ ak^i b^{i+1} / (b+1) \right] \beta \]

\[ \ldots 4.2.5.3 \]

\[ \frac{\partial \log L}{\partial a} = N \beta / a - \sum_{i=1}^{N} k^{i-1} b \beta t_i (b+1) \beta / (b+1) \beta \]

\[ \ldots 4.2.5.4 \]

\[ \frac{\partial \log L}{\partial b} = N \beta \log k + \beta \sum_{i=1}^{N} \log t_i - N(\beta - 1) / (b+1) - \]

\[ \sum_{i=1}^{N} \beta \left( ak^i b^{i+1} / (b+1) \right) \beta - 1 \cdot \frac{\partial x}{\partial b} \]

where \( x = ak^i b^{i+1} / (b+1) \)

\[ \log x = \log a + b \log K + (b+1) \log t_i - \log (b+1) \]

or, \( 1 / x \frac{\partial x}{\partial b} = (\log K + \log t_i - 1 / (b+1)) \frac{\partial b}{\partial b} \)

\[ \Rightarrow \frac{\partial x}{\partial b} = (\log K + \log t_i - 1 / (b+1)) ak^i b^{i+1} / (b+1) \]

Therefore

\[ \frac{\partial \log L}{\partial b} = N \beta \log K - N(\beta - 1) / (b+1) + \beta \sum_{i=1}^{N} \log t_i - \sum_{i=1}^{N} \left[ \beta a^i k^i b^{i+1} / (b+1) \right] \beta (\log K + \log t_i - 1 / (b+1)) \]

\[ = N \beta \log K - N(\beta - 1) / (b+1) + \beta \sum_{i=1}^{N} \log t_i - \sum_{i=1}^{N} \left[ \beta a^i k^i b^{i+1} / (b+1) \right] \beta / (b+1) \beta + 1 \]

\[ \ldots 4.2.5.5 \]
Similarly we have \( \frac{\partial \log L}{\partial \beta} \)

\[
\frac{\partial \log L}{\partial \beta} = \frac{N}{\beta} + N b \log k - N \log(b+1) + (b+1) \sum_{i=1}^{N} \log t_i + N \log a
\]

Equating these last three equations to zero. One can get the estimates \( a, b \) and \( \beta \).

ii) Estimating parameters \( a, b \) when \( \beta \) is known:

We can eliminate (4.2.5.6), and putting the known value of \( \beta \) in equation (4.2.5.4) and (4.2.5.5) and equate it to zero, one can estimate \( a \) and \( b \).

4.2.6 Properties of estimators for exponential distribution \((\beta=1)\):

The estimates \( a, b \) of equation (4.2.4.13) for \( \beta = 1 \) have the following properties.

i) \( \hat{b} \) and \( k \) are related by test data:

From the \( \text{ML} \) equation, for \( \beta = 1 \), we have

\[
\hat{a} \frac{\hat{b}^N}{(\hat{b}+1)} \sum_{i=1}^{N} t_i(\hat{b}+1) = N \quad \ldots 4.2.6.1
\]

and

\[
N \log K + \sum_{i=1}^{N} \log t_i - \sum_{i=1}^{N} \hat{a} \hat{b} t_i(\hat{b}+1) \log K/(\hat{b}+1) + \sum_{i=1}^{N} \frac{\hat{a} \hat{b} t_i(\hat{b}+1) \log t_i}{(\hat{b}+1)} - \sum_{i=1}^{N} \frac{\hat{a} \hat{b} t_i(\hat{b}+1)}{(\hat{b}+1)^2} = 0 \quad \ldots 4.2.6.2
\]

From these two equation, we have
Therefore, the $\hat{b}$ can be obtained without the slope of progressive stress $k$, that is $\hat{b}$ is indirectly related to $k$ by test data.

11) Invariance of $\hat{b}$ under the scale transformation:

Let $\omega_i = \frac{t_i}{\gamma}$, where $\gamma$ is an arbitrary constant.
Substitute $\gamma \omega_i$ for $t_i$ in equation (4.2.6.3), we have

$$\frac{N}{(\hat{b}+1)} + \sum_{i=1}^{N} \log \omega_i - \frac{N}{(\hat{b}+1)} \sum_{i=1}^{N} \log \omega_i = 0$$

$$= \frac{N}{(\hat{b}+1)} - \frac{N}{(\hat{b}+1)} \sum_{i=1}^{N} \log \omega_i$$

We see that equation (4.2.6.3) and (4.2.6.4) have the same form. Equation (4.2.6.4) does not include $\gamma$. This means that if the test data are multiplied by an arbitrary constant, $\hat{b}$ does not change.

11) Uniqueness of $\hat{a}$ and $\hat{b}$:

Let $f = \sum_{i=1}^{N} \log t_i$  \[4.2.6.5\]
\[ g(x) = N \left[ \sum_{i=1}^{N} t_i^x \log t_i \right] / \left[ \sum_{i=1}^{N} t_i^x \right] - N/x \quad \ldots 4.2.6.6 \]

then,
\[ \frac{dg(x)}{dx} = N \left[ \sum_{i=1}^{N} t_i^x \log t_i - \log t^x \right] / \left[ \sum_{i=1}^{N} t_i^x \right] \]
\[ -N \left[ \sum_{i=1}^{N} t_i^x \log t_i \right]. \left[ \sum_{i=1}^{N} t_i^x \log t_i \right] / \left[ \sum_{i=1}^{N} t_i^x \right]^2 + n/x^2 \]
\[ = N \left[ \sum_{i=1}^{N} \sum_{j=1}^{N} t_i^x \left( \log t_i - \log t_j \right)^2 \right] / \left[ \sum_{i=1}^{N} t_i^x \right]^2 + n/x^2 \quad \ldots 4.2.6.7 \]

Since \( t_i > 0, \quad i = 1, 2 \ldots N \)

Therefore \( \frac{dg(x)}{dx} > 0 \)

This means that \( g(x) \) is strictly monotone increasing function. On the other hand, we know that \( \hat{b} > 0 \) and if \( \hat{b} \) is the solution of equation \( 4.2.6.3 \), then \( g(\hat{b}+1) = f \).

Therefore, with the strict monotonicity and continuity of \( g(x) \) in \([1, +\infty]\), we conclude that \( \hat{b} \) from \( 4.2.6.3 \) is unique. Also we have \( \hat{a} \) is unique.

iv) Existence of \( \hat{b} \):

Put \( x = 1 \) in \( 4.2.6.7 \) and substract \( 4.2.6.6 \)

Thus we have
\[ g(1) - f = N \left[ \sum_{i=1}^{N} t_i^x \log t_i \right] / \left[ \sum_{i=1}^{N} t_i^x \right] - N - \sum_{j=1}^{N} \log t_j \]
\[ = \left[ \sum_{i=1}^{N} \sum_{j=1}^{N} \left( \log t_i - \log t_j \right) \cdot t_i \right] - N \left[ \sum_{i=1}^{N} t_i \right] / \left[ \sum_{i=1}^{N} t_i \right] \]
If
\[ i < j \]
then,
\[ g(l) > f \]
When \( g(l) > f \), it is certain that \( g(x) > f \), that is \( g(x) \)
is monotonic and continuous in \([1, +\infty)\). Therefore when
(4.2.6.8) is true, one cannot get solution \( b > 1 \) from
(4.2.6.3). Inversely, if
\[ i < j \]
then
\[ g(l) < f \]
But
\[ t_N > t_i, \quad i = 1, 2, 3, ..., N-1 \]
and there is at least one strict inequality.
\[
\lim_{x \to +\infty} g(x) = N \log t_N > f
\]
So, according to the continuity of \( g(x) \) on \([1 + \xi, M]\), where \( \xi \)
is a small constant and \( M \) is a large constant such that
\( g(l + \xi) < f \) and \( g(M) > f \), there exists an \( x \in [1 + \xi, M] \),
which satisfies \( g(x^*) = f \) and \( b = x^* - 1 \) is the solution
of (4.2.6.3).
Thus the necessary and sufficient condition for (4.2.6.3), providing solution \( b > 0 \) is:

\[
\sum N \sum [\log t_i - \log t_j] \cdot [t_i - t_j] < N \sum t_i.
\]

If the data from the progressive stress accelerated life testing do not satisfy this condition, then the estimate is from (4.2.6.3) is negative or zero.

********
*****
CHAPTER - V

OPTIMAL DESIGN OF ACCELERATED LIFE TESTS UNDER PERIODIC INSPECTION

5.1 Introduction:

The previous studies on accelerated life testing (ALT) is assumed continuous inspection of test items. However, further reductions in testing effort and administrative convenience may be achieved by employing periodic inspection in which test items are checked only at certain points in time. The information obtained from a periodic inspection consists of the number of failures in each inspection period, resulting in 'grouped' or 'interval' data.

Optimal ALT plans have been developed by several authors under the assumption of continuous inspection, (i.e. Chernoff, H. (1962); Little, R.E. and Jebe, E.H. (1969); Mann, N.R., (1972); Meeker, W.Q. (1984); Nelson W. and Kielpinski, T.J. (1976); Nelson, W. and Meeker, W.Q. (1978)). The present investigation is an attempt to combine these interesting and important features of life tests, namely, acceleration and periodic inspection.

This chapter contains the asymptotically optimal ALT plan for the exponential distributed lifetimes under the assumptions of Type I censoring and periodic inspection. The optimality criterion adopted is the minimum variance of the estimated mean or p th quantile of the lifetime
distribution at the use stress. Two overstress levels are considered i.e. low and high stress level, and the decision variable include the low stress level and the proportion of test items to be allocated to each stress.

The maximum likelihood estimation method is used to estimate the unknown parameters in the relationship between the mean lifetime and the stress. Then, the asymptotic properties of those ML estimators are used to approximate the variance of the estimated mean or the p\(^{th}\) quantile at the use condition.

5.2 The Model:

Assume that the lifetimes of test items are independently and identically distributed as exponential. That is,

\[ f(t) = \frac{1}{\Theta} e^{-t/\Theta}, \quad t > 0 \quad \ldots \text{5.2.1} \]

The mean lifetime \( \Theta \) and the stress \( S \) are assumed to be related as

\[ \Theta = e^{( \beta_0 + \beta_1 S)} \quad \ldots \text{5.2.2} \]

over the region of interest. The relationship can be shown that the well-known inverse power law model and the Arrhenius reaction rate model are special cases of (5.2.2).

Three stress levels are considered. That is, the use stress level \( S_0 \), the low stress level \( S_1 \) and the high stress level \( S_2 \). It is assumed that \( S_0 \) and \( S_2 \) are prespecified, while \( S_1 \) is to be optimally determined.
The number of test items allocated to $S_1$ and $S_2$ are respectively given by

$$n_1 = a_1 N$$

$$n_2 = a_2 N, = (1-a_1) N$$

Where $N$ is the total number of test items given and $a_1$ is to be optimally determined. At $S_1$, $n_1$ units are to be put on test at time $0$ and run until a specified time $t_{c1}$ (i.e. Type I censoring is assumed), and inspections are conducted only at specified points in time $t_{i1}, t_{i2}, \ldots$

$$t_{i1,k(i)} = t_{c1}$$. Also, let $t_{i0} = 0$ and $t_{i1,k(i)+1} = \infty$, and define

$$x_{ij} = \text{the number of items (at stress level $s_1$) failed in } (t_{i,j-1}, t_{ij}), j = 1, 2, \ldots, k(i)+1; \quad 5.2.5$$

$$P_{ij} = \text{the probability of failure at stress level } s_1 \text{ in } (t_{i,j-1}, t_{ij}), j = 1, 2, \ldots, k(i)+1 \quad 5.2.6$$

Then, the resulting structure of the periodic inspection at $s_1$ can be described as in figure.

Probability of failures

<table>
<thead>
<tr>
<th>$P_{i1}$</th>
<th>$P_{i2}$</th>
<th>$P_{1,k(i)}$</th>
<th>$P_{1,k(i)+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{i1}$</td>
<td>$x_{i2}$</td>
<td>$x_{i,k(i)}$</td>
<td>$x_{i,k(i)+1}$</td>
</tr>
<tr>
<td>$t_{i0}$</td>
<td>$t_{i1}$</td>
<td>$t_{i2}$</td>
<td>$t_{i,k(i)}$</td>
</tr>
</tbody>
</table>

The grouped data [ $x_{ij}$, $i = 1, 2$, $j = 1, 2, \ldots, k(i)+1$] are used to estimate $\beta_0$ and $\beta_1$, in equation (5.2.2). The estimated relationship is then extrapolated to estimate
some quantities at the use condition. Of particular
interest is the logarithm of the mean lifetime at the use
condition which is defined by
\[ \mu_o = \log \Theta_o = \beta_o + \beta_1 S_o \] ...5.2.7
Note that \( t_p \), the \( p \)th quantile of the exponential
distribution at the use condition, is related to \( \mu_o \) as
follows.
\[ \gamma_p = \log t_p = \mu_o + \log [-\log (1-p)] \] ...5.2.8
Let \( \hat{\beta}_o \) and \( \hat{\beta}_1 \) be the estimators of \( \beta_o \) and \( \beta_1 \), respectively.
Then,
\[ \hat{\mu}_o = \hat{\beta}_o + \hat{\beta}_1 S_o \] ...5.2.9
\[ \hat{\gamma}_p = \hat{\mu}_o + \log [-\log (1-p)] \] ...5.2.10
The problem of optimally designing the ALT plan
under periodic inspection can now be stated as given \( N, S_0, S_2, [t_{ij}, i = 1,2, j = 1,2...k(i)] \) determine \( \alpha_1 \) and \( S_1 \)
such that the variance of \( \hat{\mu}_o \) (or, equivalently the variance
of \( \hat{\gamma}_p \)) is minimized.

5.3 Maximum likelihood estimation and optimal plans :

At \( S_1 \), let the grouped data \([x_{ij}, j = 1,2...k(i)+1]\]
are multinomially distributed with parameters \( n_1 \) and
\([P_{ij}, j = 1,2...k(i)+1]\). The likelihood function is given
by
\[ L' = \prod_{i=1}^{2} \prod_{j=1}^{k(i)+1} \frac{n_1!}{x_{ij}!} \left( \prod_{j=1}^{k(i)+1} \frac{P_{ij} x_{ij}}{n_1} \right) \] ...5.3.1
Taking logarithm both side, we have

\[ L = \log L' = \sum_{i=1}^{2} \log L_i' \]

\[ = C + \sum_{i=1}^{2} \log (\sum_{j=1}^{k(i)+1} p_{ij} x_{ij}) \]

\[ = C + \sum_{i=1}^{2} \sum_{j=1}^{k(i)+1} x_{ij} \log p_{ij} \quad \ldots 5.3.2 \]

where C is constant

Note that

\[ p_{ij} = e^{-(t_{ij} - 1)/\Theta_1} \cdot e^{-(t_{ij} - 1)/\Theta_1} \quad \ldots 5.3.3 \]

for \( i = 1,2 \) and \( j = 1,2 \ldots k(i)+1 \)

Then the likelihood function becomes

\[ L = C + \sum_{i=1}^{2} \sum_{j=1}^{k(i)+1} x_{ij} \log [e^{-(t_{ij} - 1)/\Theta_1} - e^{-(t_{ij} - 1)/\Theta_1}] \]

To find ML estimates of \( \beta_0 \) and \( \beta_1 \), we differentiate the above equation with respect to \( \beta_0 \) and \( \beta_1 \), we have

\[ \frac{\partial L}{\partial \beta_0} = 2 \sum_{i=1}^{2} \sum_{j=1}^{k(i)+1} x_{ij} \frac{1}{p_{ij}} \left[ t_{ij} - \frac{1}{\Theta_1} \cdot e^{-(t_{ij} - 1)/\Theta_1} - \frac{t_{ij}}{\Theta_1} \cdot e^{-(t_{ij} - 1)/\Theta_1} \right] \]

Since \( \Theta_1 = e^{\beta_0} + \beta_1 s_i \)

\[ = 2 \sum_{i=1}^{2} \sum_{j=1}^{k(i)+1} x_{ij} \frac{(B_i, j-1 - B_{ij})}{p_{ij}} \quad \ldots 5.3.4 \]

Similarly we have
\[
\frac{\partial L}{\partial \beta_1} = \sum_{i=1}^{k(1)+1} x_{ij} \left( B_{i,j-1} - B_{i,j} \right) / P_{ij} \quad \text{...5.3.5}
\]

where \( B_{i,j} = t_{ij}/\theta_1 \), \( e^{-t_{ij}/\theta_1} \) \text{...5.3.6}

Put these equation equal to zero and after solving, one can get the estimate of \( \beta_0 \) and \( \beta_1 \).

The Fisher information matrix is given by

\[
F = N \left( f_{gh} \right) g, h = 0, 1, \quad \text{...5.3.7}
\]

\[
f_{gh} = \sum_{i=1}^{k(1)+1} a_i \sum_{j=1}^{s_i} \left( \frac{\partial P_{ij}/\partial \beta_g}{P_{ij}} \frac{\partial P_{ij}/\partial \beta_h}{P_{ij}} \right) \quad \text{...5.3.8}
\]

After some algebraic manipulation, we obtain

\[
f_{oo} = \sum_{i=1}^{2} a_i \theta_i \quad \text{...5.3.9}
\]

\[
f_{ol} = f_{lo} = \sum_{i=1}^{2} a_i S_i \theta_i \quad \text{...5.3.10}
\]

\[
f_{ll} = \sum_{i=1}^{2} a_i S_i^2 \theta_i \quad \text{...5.3.11}
\]

where

\[
Q_i = \sum_{j=1}^{k(1)+1} \left( B_{i,j-1} - B_{i,j} \right)^2 / P_{ij} \quad \text{...5.3.12}
\]

Then, the asymptotic covariance matrix of \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) is obtained as

\[
V = F^{-1} = N^{-1} \begin{pmatrix} v_{oo} & v_{ol} \\ v_{lo} & v_{ll} \end{pmatrix}
\]

\[
= N^{-1} \left( f_{oo} f_{ll} - f_{ol}^2 \right)^{-1} \begin{pmatrix} f_{ll} & -f_{ol} \\ -f_{lo} & f_{oo} \end{pmatrix} \quad \text{...6.3.13}
\]

From equation (5.2.9), the asymptotic variance (avar) of \( \hat{\mu}_0 \) is given by
avar (µ_) = avar (β_0 + β_1 S_o)

= N^{-1} (v_{00} + s_o^2 v_{ll} + 2 s_o v_{0l})

= N^{-1} (f_{00} f_{ll} - f_{0l}^2)^{-1} (f_{ll} + s_o^2 f_{00} - 2 s_o f_{0l}) \ldots 5.3.14

which is also avar (γ_p)

The optimization problem is to determine S_1 and α_1 that minimize avar (µ_0). We first assume without loss of generality that the stress level is adjusted such that S_o = 0, then eqn. (5.3.14) is reduced to

avar (µ_0) = avar (β_0)

= N^{-1} (f_{00} f_{ll} - f_{0l}^2)^{-1} f_{ll} \ldots 5.3.15

Substituting the value, we obtain

= N^{-1} \left[ \sum_{i=1}^{2} \alpha_i Q_i \left( \sum_{i=1}^{2} \alpha_i S_i Q_i - \sum_{i=1}^{2} \alpha_i S_i Q_i \right) \right]^{-1} \frac{2}{\sum_{i=1}^{2} \alpha_i S_i Q_i}

= N^{-1} \left[ (\alpha_1 Q_1 + \alpha_2 Q_2) (\alpha_1 S_1^2 Q_1 + \alpha_2 S_2^2 Q_2) - (\alpha_1 S_1 Q_1 + \alpha_2 S_2 Q_2)^2 \right]^{-1}

\left( \alpha_1 S_1^2 Q_1 + \alpha_2 S_2^2 Q_2 \right)

= N^{-1} \left[ \alpha_1 \alpha_2 Q_1 Q_2 S_2^2 + \alpha_1 \alpha_2 Q_1 Q_2 S_1^2 - 2 \alpha_1 \alpha_2 Q_1 Q_2 S_1 S_2 \right]^{-1}

\left( \alpha_1 S_1^2 Q_1 + \alpha_2 S_2^2 Q_2 \right)

= N^{-1} \left[ \alpha_1 S_1^2 Q_1 + \alpha_2 S_2^2 Q_2 \right]/ \left[ \alpha_1 \alpha_2 Q_1 Q_2 (S_1 - S_2)^2 \right]

= N^{-1} \left[ \alpha_1 S_1^2 Q_1 + (1-\alpha_1) S_2^2 Q_2 \right]/ \left[ \alpha_1 (1-\alpha_1) Q_1 Q_2 (S_1 - S_2)^2 \right]

= N^{-1} \left[ S_1^2 Q_1 + S_2^2 Q_2 \right]/ \left[ Q_1 Q_2 (S_1 - S_2)^2 (-\alpha_1^2 + \alpha_1) \right]
The following two step procedure is obtained to minimize\( \text{avar}(\hat{\mu}_0) \) with respect to \( S_1 \) and \( \alpha_1 \).

First, we optimize \( \alpha_1 \) for given \( S_2 \). The partial derivative of \( \text{avar}(\hat{\mu}_0) \) with respect to \( \alpha_1 \) is given by

\[
\frac{\delta(\text{avar}(\hat{\mu}_0))}{\delta \alpha_1} = N^{-1} \left[ (S_1^2 Q_1 - S_2^2 Q_2) \alpha_1^2 + 2 S_1^2 Q_2 \alpha_1 - S_2^2 Q_2 \right] / [Q_1 Q_2 (S_1 - S_2)^2 (-\alpha_1^2 + \alpha_1)^2]
\]  

(5.3.17)

Let \( f(\alpha_1) \) be the term in the first braces of equation (5.3.17). Then, \( f(\alpha_1) \) behaves as shown in figure (2). The sign of \( \frac{\delta(\text{avar}(\hat{\mu}_0))}{\delta \alpha_1} \) is always the same as \( f(\alpha_1) \) because \( N \) and the term in the second braces of equation (5.3.17) are always positive. Therefore, for \( 0 < \alpha_1 < 1 \), \( \text{avar}(\hat{\mu}_0) \) behaves as Figure 3, where \( \alpha_1^* \), the optimal value of \( \alpha_1 \), is given by

\[
\alpha_1^* = (-S_2^2 Q_2 + \sqrt{S_1^2 S_2^2 Q_1 Q_2} / (S_1^2 Q_1 - S_2^2 Q_2))
\]  

(5.3.18)

The second stage of optimization employs the grid search with respect to \( S_1 \). That is \( \alpha_1^* \) and the minimum of \( \text{avar}(\hat{\mu}_0) \) are determined on the grid \( S_1 = d, 2d, 3d, \ldots \), where \( d \) is the grid size. Final optimal solution consists of \( S_1^* \) (and the corresponding \( \alpha_1^* \)) for which \( \text{avar}(\hat{\mu}_0) \) attains the minimum among all the grid points considered.
Figure 2. The behavior of $f(a_i)$.

Figure 3. The behavior of $\text{avar}(\hat{\mu}_i)$ for given $s_i$. 
When \( k(1) \) and \( k(2) \) go to infinite (i.e., continuous inspection) it can be shown that \( \text{avar}(\hat{\mu}_0) \) has the same form as equation (5.3.16) with

\[
Q_1 = 1 - e^{-\frac{t_{ci}}{\theta_1}}
\]

Therefore, \( S_1^* \) and \( a_1^* \) for continuous inspection can be also determined by the above two-step procedure.
REFERENCES


***