

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA**

UMI[®]
800-521-0600

Accelerated Life Testing: Concepts and Models

Debaraj Sen

**A Thesis
in
The Department
of
Mathematics and Statistics**

**Presented in partial fulfillment of the requirements
for the degree of Master of Science at
Concordia University
Montreal, Quebec, Canada**

August 1999
© Debaraj Sen, 1999



**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

395 Wellington Street
Ottawa ON K1A 0N4
Canada

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-43638-1

Canada

Abstract

Accelerated Life Testing: Concepts and Models

Debaraj Sen

This thesis deals with the analysis of accelerated life test. First, we provide related concepts and second, we provide detailed properties of three alternative distributions, Gamma, Log-normal and Inverse Gaussian. Various models for failure times are considered which are plausible in this context and estimation procedure using the ML method is outlined. Finally, a numerical example is considered using various models introduced earlier.

Acknowledgement

I am grateful to my supervisor Prof. Yogendra P. Chaubey for accepting me as his student and suggesting to work in this field. Throughout the preparation of this thesis, his continuous encouragement, constructive discussions, important suggestions and necessary corrections of my work have made the success of this work possible. I also acknowledge his hearty cooperation in providing me with all his valuable reference materials on this area.

I am thankful to Prof. J. Garrido and Prof. T.N. Srivastava for agreeing to be on my thesis committee.

I am grateful to the Department of Mathematics and Statistics at Concordia University for its financial support as well as every Professor and Secretary who helped me to overcome any technical or theoretical problem.

I wish to express my earnest thanks to all my friends who encouraged and inspired me in many ways and I also thank to Mr. Anthony Crisalli for his valuable suggestions.

I am grateful to my parents for making my entire education possible.

Contents

1. Accelerated life testing and related concepts	
1.1 Introduction	1
1.2 Stress and its classification	2
1.3 Type of censoring	4
1.4 Relevance of statistical methods	5
2. Accelerated life models based on Transferable functions	
2.1 Transferable functions	6
2.2 Accelerated life models	7
2.3 Survival functions in the case of the models	10
2.4 Accelerated life models when the stress is constant	11
2.5 Accelerated life models when the stress is piecewise constant	15
2.6 Accelerated life models when the stress is progressive	19
3. Possible survival distributions and their properties	
3.1 Introduction	21
3.2 Gamma distribution	
3.2.1 Gamma distribution as a model for failure times	21
3.2.2 Properties	22
3.2.3 Applications and uses	24
3.3 Log-normal distribution	
3.3.1 Introduction	25
3.3.2 Properties	25
3.3.3 Applications and uses	26
3.4 Inverse Gaussian distribution	
3.4.1 Introduction	27
3.4.2 Properties	28
3.4.3 Applications and uses	29
4. Models of failure times in accelerated life testing	
4.1 The reciprocal linear regression model	31
4.2 Introduction to ML method	32
4.3 Computational aspects of ML method	33
4.4 Asymptotic properties of maximum likelihood estimators	34
4.5 The Newton-Raphson method	36
4.6 Convergence of the Newton-Raphson method	39
4.7 The Scoring method	40
4.8 Review of accelerated life testing	
4.8.1 Nelson's work	42
4.8.2 Work of Bhattacharyya and Fries	43
4.8.3 Work of Singpurwalla	44
4.8.4 Work of Babu and Chaubey	45

4.9 Estimation of the parameters of the reciprocal linear regression model by maximum likelihood method	45
5. Numerical illustration	
5.1 Data	49
5.2 Estimates of the parameters	49
5.3 Goodness of fit of the models	52
5.4 Conclusions and future research	53
Appendix	55
Bibliography	63

1 Accelerated Life Testing and Related

Concepts

1.1 Introduction

The term “Accelerated life test” applies to the type of study where failure times can be accelerated by applying higher “stress” to the component. This implies that the failure time is a function of the so called “stress factor” and higher stress may bring quicker failure. For example, some component may fail quicker at a higher temperature however, it may have a long life at lower temperatures. At low ‘stress’ conditions, the time required may be too large for its reliability estimation which may be tested under higher stress factors terminating the experiment in a relatively shorter time. By this process failures which under normal conditions would occur only after a long testing can be observed quicker and the size of data can be increased without a large cost and long time. This type of reliability testing is called “Accelerated life testing”.

Accelerated life testing methods are also useful for obtaining information on the life of products or materials over a range of conditions, which are encountered in practice. Some information can be obtained by testing over the range of conditions of interest or over more severe conditions and then extrapolating the results over the range of interest. This type of test conditions are typically produced by testing units at high levels of temperature, voltage, pressure, vibration, cyclic rate, load etc. or some combination of them. Stress variables are used in engineering practice for many products and materials. In other fields similar problems arise when the relationship between variables could affect its life time. Therefore the models formulated are based on either past studies or theoretical development that could relate the distribution of failure time to stress or other variables. Such models are also useful in survival analysis where dependence of the life time of individuals on concomitant variables is analyzed.

The idea of life testing is briefly discussed by Nelson [65] and in a report by Yurkowsky *et al.*[106]. The bibliography by Goba [37] gives a list of categorized references on thermal aging of electrical insulation. The general area of accelerated life testing deals with the testing methods, model consideration, form of the life data and required statistical methods. In the next few sections, we will discuss the details of various aspects involved in accelerated life testing. Section-2 presents the generalized concept of stress and its classifications. Section-3 presents the

censoring aspect of the data and section-4 presents the relevance of statistical methods.

The purpose of this thesis is to provide an overview of statistical models and methods used along with application of these models on some real life data. Chapter 2 gives details of accelerated life models based on transferable functions where as Chapter 3 describes the details of three life time distributions namely Gamma, Log-normal and Inverse Gaussian distributions. These distributions are used later in numerical computations. Chapter 4 describes the reciprocal linear regression model and other related models and also estimate their parameters.

1.2 Stress and Its Classification

Let us consider a non-negative random variable $T(x)$ which represents the time of failure of an item depending on a vector of covariates x . In reliability theory, the vector x is called a stress vector. The probability of failure of an item is then a function of stress given by

$$F_x(t) = P[T(x) > t], t \geq 0 .$$

Assume that $F_x(t)$ is differentiable and decreasing on $(0, \infty)$ as a function of t for every $x \in S$, where S denotes a set of possible stress values.

In accelerated life testing, stress is classified into constant stress, step stress, progressive stress, cyclic stress and random stress. We now define and describe these stress classifications (see Nelson [1990]).

(a) Constant stress

In constant stress testing, each test unit is observed until it fails, keeping all the stress factors at constant levels. For some materials and products, accelerated test models for constant stress are better developed and experimentally established. Examples of constant stress are temperature, voltage and current.

(b) Step stress

In step stress, a specimen is subjected to successively higher levels of stress. At first, it is subjected to a specified constant stress for a specified length of time. If it does not fail, it is subjected to a higher stress level for a specified time. The stress on a unit is increased step by step until it fails. Usually all specimens go through the same specific pattern of stress levels and test times. But sometimes different patterns are applied to different specimens. The increasing stress levels ensure that the failures occur quickly resulting in data appropriate for statistical

purposes. The problem with this process is that most products run at constant stress in practice not in step stress. So the model must take properly into an account of the cumulative effect of exposure at successive stresses and it must provide an estimate of life under constant stress. In a step stress test, the failure models occur at high stress levels these may differ from normal stress conditions. Such a test yields no greater accuracy than a constant stress test of the same length.

(c) Progressive stress

In this type of stress, a specimen undergoes a continuously increasing level of stress. This test is also called a port test and it is used to determine the endurance limit of a metal in the study of metal fatigue. These stress tests also have the same disadvantage as the step-stress test. Moreover, it may be difficult to control the accuracy of the progressive stress.

(d) Cyclic stress

A cyclic stress test repeatedly loads a specimen with the same stress pattern. For many products, a cycle is sinusoidal. For others, the test cycle repeats but is not sinusoidal. For insulation tests, the stress level is the amplitude of the AC voltage . Therefore a single number characterizes the level. But in metal fatigue tests, two numbers characterize such sinusoidal loading. Thus, fatigue life can be regarded as a function of these two constant stress variables. In most cases, the frequency and length of a stress cycle are the same as in actual product use. But sometimes, they are different and may be assumed to have negligible effect on the life time of the product and therefore they are disregarded. For many products, the frequency and length of a cycle affect the life time of the product, so they are included in the model as a stress variable.

(e) Random stress

In random stress testing, some products are used to undergo randomly changing levels of stress. For example, bridge and airplane structural components undergo wind buffeting. Also environmental stress screening uses random vibration. So an accelerated test employs random stresses but at higher levels.

In survival analysis, the covariate x is a vector, components of which correspond to various characteristics of life time of individuals such as methods of cure operation, quantities or types of remedies, environment, interior characteristics such as blood pressure, sex. These factors can be constant or non-constant in time.

Among all stresses, the constant stress method has been used widely and it is considered more important than other stress testing methods. The relationship between the accelerating variable

and life time are well developed in this case. Because this method requires long test times, step stress and progressive stress testing which require shorter test times, may be employed. However in progressive stress testing, it is difficult to maintain a constant rate of increase so that in this situation step stress testing method is easier to carry-out. Actually, the relation between the accelerating variables and life time depends on the pattern of step or progressive testing. It requires a model for the cumulative damage, *i.e.* the result on life of the exposure changes its environment. Such models are more difficult than those for constant stress. These type of testing methods are used in an experiment with one accelerating variable. More than one accelerating variable are discussed in Nelson [65] and they have involved only constant stress testing.

1.3 Type of Censoring

An important reason for special statistical models and methods for failure time data is to accommodate the right censoring in the data. Generally censoring complicates the distribution theory for the estimator even when the censoring technique is simple and in other cases complex censoring technique may make some computations impossible. If the values of the observation in one of the distribution tails is unknown then the data is said to be single censored. For example, life test data are single censored because in a life test if all units are placed on test at the same time and all unfailed units have accumulated the same running time at the time of analysis, the failure times of unfailed units are known only beyond their current running times. If the values of the observations in both tails are unknown then data are called double censored. For example, instrumentation data may be doubly censored because observations may be beyond the scale of measurement at either tails of the distribution. Some data censored on the right have differing running times inter-mixed with the failure times. These type of data are called multiply or progressively censored. Censored data are called type-I censored if observations occur only at specified values of dependent variable. For example, in life testing when all units are put on test at the same time and the data are collected and analyzed at a particular point in time. For life data, it is also called time censored if the censoring time is fixed and the number of failures in the fixed time is random. But if the number of censored observations is specified and their censored values are random then this type of censored data are type-II censored. For life data, it is called failure censored if the test is stopped when a specified number of failures occur and the

time to that fixed number of failures is random.

In accelerated life testing, data are analyzed before all specimens fail. If the model and data are valid then the estimates from the censored data are less accurate than those from the complete data.

1.4 Relevance of Statistical Methods

Accelerated life tests serve various purposes. *i.e.*, identify design failures, estimate the reliability improvement by eliminating certain failure models, determine burn in time and conditions, quality control, assess whether to release a design to manufacturing or product to a customer, demonstrate product reliability for customer specifications, determine the consistency of the engineering relationship and adequacy of statistical models, develop relationship between reliability and operating conditions. Actually, management must specify accurate estimates for their design purposes and statistical test planning helps towards this goal. Many engineering experiments may not follow a good experimental design, without which analysis and interpretations of data may not be adequate and thus may result in improper decision.

2 Accelerated Life Models Based on Transferable Functions

2.1 Transferable Functions

Let $T(x)$ be a non-negative random variable representing the time of failure which is considered to depend on a vector of covariates x and $R_x(t)$ be the survival function of the random variable $T(x)$ given by

$$R_x(t) = P\{T(x) > t\} = 1 - F_x(t), t \geq 0.$$

Let x_0 be the base line stress level, also known as the normal stress in reliability theory. Then the function f is defined by

$$f(t, x) = F_{x_0}^{-1}[F_x(t)], t > 0 \quad (2.1.1)$$

and is called the transferable function. We assume here that $F_x(t)$ is absolutely continuous for every x . Equation (2.1.1) implies that

$$P\{T(x) \leq t\} = P\{T(x_0) \leq f(t, x)\}. \quad (2.1.2)$$

i.e. the probability that an item under stress x would survive at time t is equal to the probability that an item used under normal stress x_0 would survive beyond $f(t, x)$. The time t under any stress x is in this sense equivalent to $f(t, x)$ under the base line stress x_0 , then $f(t, x)$ is called the resource used at time t under stress x . The random variable $R(x) = f(T(x), x)$ is called simply the resource under the failure time $T(x)$ and the survival function of $R(x)$ is R_{x_0} .

Again suppose G is some survival function defined on $[0, \infty)$ and there exists an inverse function $H = G^{-1}$, the function defined by

$$f(t, x) = H[R(x)(t)], \quad (2.1.3)$$

is called the G -transferable function or simply transferable function and we have

$$P\{T(x) \leq t\} = P\{T_G \leq f(t, x)\}, \quad (2.1.4)$$

where T_G is a random variable having the survival function G and the corresponding resource known as the G resource is given by

$$R = (H \cdot R(x))T(x),$$

which has survival function G .

2.2 Accelerated Life Models

The models of accelerated life testing can be formulated on the basis of the properties of the transferable function. To obtain a larger class of models we generalize the notion of the transferable function rejecting the assumption that the distribution of the resource is defined by the survival function. Consider E to be the subset of some set of stresses.

Model-1

This model is said to hold if there exists a positive functional $r : E \rightarrow (0, \infty)$ such that for all $x \in E$ the transferable function f satisfies the differential equation

$$\frac{df(t, x)}{dt} = r(x(t)),$$

subject to the condition that $f(0, x) = 0$. The model implies that the rate of change of the resource depends only on the covariate x at the time t . Now

$$f(t, x) = \int_0^t r(x(u))du. \quad (2.2.1)$$

Therefore the resource is

$$R = \int_0^{T(x)} r(x(u))du. \quad (2.2.2)$$

In this model, the covariate changes only the scale of the time to failure distribution. This model is known as accelerated failure time model (see Nelson [65]).

Model-2

This model is said to hold if there exists a positive functional $r : E \rightarrow (0, \infty)$ such that for all $x, y \in E$ the transferable function f satisfies the differential equation

$$\frac{df(t, x)}{dt} \div \frac{df(t, y)}{dt} = r(x(t)) \div r(y(t)),$$

subject to the condition that $f(0,x) = f(0,y) = 0$. This model implies that the ratio of rates of resource at the two points x and y depends only on values x and y at the time t . If the covariates are fixed i.e, $y = x_0 \in E$, denote $r(x(T)) \div r(y(T))$ by $r(x(T))$, then for all $x \in E$ we have

$$f(t,x) = \int_0^t r(x(u))df(u,x_0) \quad (2.2.3)$$

and the resource is

$$R = \int_0^{T(x)} r(x(T))df(T,x_0). \quad (2.2.4)$$

Here the distribution of R does not depend on $x \in E$. This model is the generalization of the general transformation model (see Dabrowska [26]) which is formulated using an unknown monotone function h such that $h(T(x)) = \beta^T x + \epsilon$ where ϵ follows some known distribution and $\beta = (\beta_1, \beta_2, \dots, \beta_m)^T$ is a vector of unknown parameters.

Model-3

This model is said to hold if there exists a positive functional $r : E \rightarrow (0, \infty)$ such that for all $x, y \in E$ the transferable function satisfy the differential equation

$$\frac{d}{dt}f(t,x) = \frac{d}{dt}f(t,y) + a(x(t)) - a(y(t)),$$

subject to the condition that $f(0,x) = f(0,y) = 0$. This model means that the difference of rates of resource depends only on the values of the covariates x and y at the time t . For some fixed covariate $y = x_0 \in E$ denote

$$a(x(u)) = a(x(u)) - a(x_0(u)),$$

then for all $x \in E$ the function is

$$f(t,x) = f(t,x_0) + \int_0^t a(x(u))du.$$

Therefore the resource is

$$R = f(T(x),x_0) + \int_0^{T(x)} a(x(u))du. \quad (2.2.5)$$

Model 1, 2 and 3 can be considered as parametric when the survival function G is supposed to be

from some parametric family of distributions and the functions have some specified forms depending on some unknown parameters.

Model-4

This model is said to hold if there exists a positive functional $r : E \rightarrow (0, \infty)$ such that for all $x \in E$ the transferable function f satisfies the differential equation

$$\frac{d}{dt}f(t, x) = r[x(t)]q[f(t, x)],$$

subject to the condition that $f(0, x) = 0$. Here q is some positive function on x . This model means that the rate of resource is proportional to some functional of the stress at the time t . Therefore the transferable function is

$$f(t, x) = \int_0^t r[x(u)]q[f(u, x)]du, \quad (2.2.6)$$

and the resource is

$$R = \int_0^{T(x)} r[x(u)]q[f(u, x)]du. \quad (2.2.7)$$

Model-5

This model is said to hold if there exists a positive functional $r : E \rightarrow (0, \infty)$ such that for all $x \in E$ the transferable function f satisfy the differential equation

$$\frac{d}{dt}f(t, x) = g(x(t), f(t, x)),$$

subject to the condition that $f(0, x) = 0$. Here g is some positive function on $[0, \infty)$. This model means that the rate of resource depends on the value of a covariate at the time t . So the transfer function is

$$f(t, x) = \int_0^t g(x(u), f(u, x))du, \quad (2.2.8)$$

and the resource is

$$R = \int_0^{T(x)} g(x(u), f(u, x))du. \quad (2.2.9)$$

2.3 Survival Functions in the Case of the Models

For all $x \in E$ the survival functions of the model 1, 2 and 3 have the following forms

$$\left. \begin{aligned} R(x(t)) &= G\left\{\int_0^t r(x(u))du\right\} \\ R(x(t)) &= G\left\{\int_0^t r(x(u))dH(R_0(u))\right\} \\ R(x(t)) &= G\left\{H(R_0(u)) + \int_0^t a(x(u))du\right\} \end{aligned} \right\} \quad (2.3.1)$$

Here R_0 is some fixed survival function but it does not depend on $x \in E$. But in the case of models 4 and 5 the survival function $R(x)$ satisfies the equation

$$\left. \begin{aligned} H(R(x(t))) &= \int_0^T r(x(u))q_0(R(x(t)))du \\ H(R(x(t))) &= \int_0^T g_0[x(u), R(x(t))]du \end{aligned} \right\} \quad (2.3.2)$$

Now consider the class of constant in time covariates then the survival function of the models 1,2 and 3 have the following forms

$$\left. \begin{aligned} R(x(t)) &= G\{r(x)t\} \\ R(x(t)) &= G\{r(x)H(R_0(t))\} \\ R(x(t)) &= G\{H(R_0(t)) + q(x)t\} \end{aligned} \right\} \quad (2.3.3)$$

The models 1,2 and 3 are parametric if the function G is supposed to be from some parametric family of survival functions and r and q have some specified forms depends on some unknown parameters. If the function R_0 is completely unknown and the functions G and (or) r are parametrized then the models are semi-parametric. But in model-2, the distribution G can often be parameter free because this model is invariant with respect to some transformation. If the functions R_0, r and G are completely unknown then the models are non-parametric.

2.4 Accelerated Life Models when the Stress is Constant

Consider the sub-models of the models 2 and 3 for the class of constant in time stresses.

Model-2

Suppose that model 2 is holds for all $x \in E_0$ then

$$R(x(t)) = G\{r(x)H(R_0(t))\}, \quad (2.4.1)$$

where R_0 is some unknown survival function. But it is invariant with respect to the transformation $S = \left(\frac{t}{\theta}\right)^\alpha$ when $G(t) = G_0\left(\left(\frac{t}{\theta}\right)^\alpha\right)$ for $\theta > 0$, $\alpha > 0$. Thus all the distributions of the resource from the class $G_0\left(\left(\frac{t}{\theta}\right)^\alpha\right)$ follow the same model and $\theta = \alpha = 1$ always can be taken. In this case, the important idea for accelerated experiment is that the accelerating factor is one dimensional and the parametrization is $r(x) = a \exp(bz)$ where $z = z(x)$ is some unknown function.

Consider the model, $R(x(t)) = G\{ae^{\beta x}H(R_0(t))\}$ for some function of the stress or the stress x . Suppose the reliability function R is

$$R(t) = G\{aH(R_0(t))\}.$$

Then we obtain

$$G\{ae^{\beta x}H(R_0(t))\} = G\{e^{\beta x}H(R(t))\},$$

therefore the new model is

$$R(t) = G\{e^{\beta x}H(R(t))\}. \quad (2.4.2)$$

Now if the reliability function G has the form $G = G_0\left(\left(\frac{t}{\theta}\right)^\alpha\right)$ then there is only one parameter β to be estimated. Below we consider some specific distributions of the resource:

(a) Gamma resource

The density of the gamma distribution is

$$g(t, m, \theta) = \frac{1}{\theta^m \Gamma(m)} e^{-\frac{t}{\theta}} t^{m-1} I_{(0, \infty)}(t) \text{ for } m, \theta > 0. \quad (2.4.3)$$

The survival function of the resource is given by

$$G(t, m, \theta) = \frac{1}{\Gamma(m)} \int_{\frac{t}{\theta}}^{\infty} x^{m-1} e^{-x} dx I_{(0, \infty)}(t). \quad (2.4.4)$$

Taking parameter free survival function, we have

$$G(t, m) = \int_t^{\infty} x^{m-1} e^{-x} dx$$

and the formula (2.4.1) implies the model

$$I(R(x(t)), m) = r(x) I(R_0(t), m), \quad (2.4.5)$$

where $I(R_0(t), m)$ is the incomplete gamma distribution with the survival function $G(t, m)$. If $m = 1$ this model is the famous Cox proportional hazard model.

(b) Weibull resource

The density function of Weibull distribution is

$$g(t, \alpha, \theta) = c \alpha^{-1} \left\{ \frac{(t-\theta)}{\alpha} \right\}^{c-1} \exp \left[- \left\{ \frac{(t-\theta)}{\alpha} \right\}^c \right] I_{(0, \infty)}(t),$$

with $\theta > 0$, $\alpha > 0$, $t > \theta$ and the survival function is

$$G(t, \alpha, \theta) = \exp \left[- \left\{ \frac{t}{\theta} \right\}^\alpha \right] I_{(0, \infty)}(t) \text{ for } \alpha > 0, \theta > 0. \quad (2.4.6)$$

Now we adopt a parameter free survival function of the resource

$$G(t) = e^{-t} I_{(0, \infty)}(t),$$

and the formula (2.4.1) implies the model

$$R[x(t)] = (R_0(t))^{r(x)} \quad (2.4.7)$$

$$R'[x(t)] = r(x) (R_0(t))^{r(x)-1} \quad (2.4.8)$$

Thus the equality (2.4.1) implies for the failure rate of random variable

$$\lambda_x(t) = \frac{-R'[x(t)]}{R[x(t)]} = r(x) \lambda_0(t). \quad (2.4.9)$$

It is the proportional hazard model (See Bagdonavičius and Nikulin [8]).

(c) log-normal resource

The density function of the log-normal distribution is

$$g(t, \alpha, \theta) = \left[(t - \theta) \sqrt{2\pi} \delta \right]^{-1} \exp \left[-\frac{1}{2} \frac{\{\log(x - \theta) - \alpha\}^2}{\sigma^2} \right], \quad (2.4.10)$$

with $\alpha > 0$, $\theta > 0$ and the survival function is

$$G(t, \alpha, \theta) = \Phi \left(\log \left(\frac{t}{\theta} \right)^\alpha \right) I_{(0, \infty)}(t) \quad \text{for } \alpha > 0, \theta > 0. \quad (2.4.11)$$

Now we adopt a parameter free survival function for $t > 0$ which is

$$G(t) = \Phi(\log(t)) I_{(0, \infty)}(t),$$

Therefore formula (2.4.1) implies that

$$\Phi^{-1}(R[x(t)]) = \log r(x) + \Phi^{-1}(R_0(t)). \quad (2.4.12)$$

It is the generalized probit model (see Dabrowska and Doksum [26]).

(d) log logistic resource

The density of the log- logistic distribution is

$$g(t, \alpha, \theta) = \alpha \theta (\alpha t)^{\theta-1} (1 + (\alpha t)^\theta)^{-2} \quad \text{for } \alpha > 0, \theta > 0 \quad (2.4.13)$$

The survival function is

$$G(t, \alpha, \theta) = \left(1 + \left(\frac{t}{\theta} \right)^\alpha \right)^{-1} I_{(0, \infty)}(t) \quad \text{for } \alpha > 0, \theta > 0.$$

So we adopt a parameter free survival function which is

$$G(t) = \frac{1}{1+t} I_{(0, \infty)}(t).$$

The formula (2.4.1) implies the model

$$\log \frac{R[x(t)]}{1 - R[x(t)]} = -\log r(x) + \log \frac{R_0[x(t)]}{1 - R_0[x(t)]}. \quad (2.4.14)$$

This is the analogue of the model for logistic regression which is used for analysis of dichotomous data when the probability of success or failure is dependent of some factors (see Lawless [57]). We can consider other distributions for the resource. If larger classes of

distributions of the resource are considered then the more general semi-parametric models including known models can be obtained.

Model-3

For all $x \in E_0$, the model is

$$R(x(t)) = G\{H(R_0(t)) + q(x)t\}. \quad (2.4.15)$$

Here R_0 is some unknown survival function. For $\theta > 0$ if $G(t) = G_0\left(\frac{t}{\theta}\right)$ then the equation (2.4.15) is invariant with respect to the transformation $S = \frac{t}{\theta}$ and $\theta = 1$ can always be taken. The important idea for accelerating experiment is when the accelerating factor is unidimensional then the parametrization of $q(x)$ is $q(x) = \alpha + \beta z$ where $z = z(x)$ is some known function. But always we will take $q(x) = \beta z$. From equation (2.4.15) we have

$$R(x(t)) = G\{H(R_0(t)) + (\alpha + \beta z)t\}. \quad (2.4.16)$$

Suppose

$$R(t) = G\{H(R_0(t)) + \alpha t\},$$

Then we obtain

$$G\{H(R_0(t)) + (\alpha + \beta z)t\} = G\{H(R(t)) + \beta zt\}.$$

So the new model is

$$R(x(t)) = G\{H(R(t)) + \beta zt\}. \quad (2.4.17)$$

which can be considered. More generally, if $q(x) = \alpha + \beta^T z$ where $\beta = (\beta_1, \dots, \beta_m)^T$ and $z = (z_1, \dots, z_m)^T$ and β is the vector of regression parameters and z is a vector of some unknown stress functions then the parameter α can be eliminated and this model is considered as semi-parametric.

Consider some specified distributions of the resource.

(a) Weibull resource

The density function of Weibull distribution is

$$g(t, \alpha, \theta) = c\alpha^{-1} \left\{ \frac{(t-\theta)}{\alpha} \right\}^{c-1} \exp \left[- \left(\frac{(t-\theta)}{\alpha} \right)^c \right] I_{(0,\infty)}(t) \text{ for } \theta > 0, \alpha > 0$$

and the survival function is

$$G(t) = e^{-t^\alpha}.$$

So model-3 simplifies to

$$\Lambda_x^{\frac{1}{\alpha}}(t) = \Lambda_0^{\frac{1}{\alpha}}(t) + q(x)t. \quad (2.4.18)$$

If the resource is exponential, i.e. $\alpha = 1$ then the additive risk model is obtained, i.e.

$$\lambda_x(t) = \lambda_0(t) + q(x). \quad (2.4.19)$$

(b) log-logistic resource

The survival function of this distribution is

$$G(t) = \frac{1}{(1+t^\alpha)}. \quad (2.4.20)$$

so model-3 implies

$$\left(\frac{1 - R(x(t))}{R(x(t))} \right)^{\frac{1}{\alpha}} = \left[\frac{1 - R_0(t)}{R_0(t)} \right]^{\frac{1}{\alpha}} + q(x)t. \quad (2.4.21)$$

(c) lognormal resource

The survival function of this distribution is

$$G(t) = \Phi(\alpha \log t), \quad (2.4.22)$$

where Φ is the survival function of the standard normal distribution. Therefore the model simplifies to

$$e^{\frac{1}{\alpha}\Phi^{-1}(R(x(t)))} = e^{\frac{1}{\alpha}\Phi^{-1}(R_0(t))} + q(x)t. \quad (2.4.23)$$

2.5 Accelerated Life Models when the Stress is Piecewise Constant

Consider the possible submodels for models 1 to 5 for the class $E_1 \subset E$ of stepwise stresses is of the form

$$x(u) = x_i \text{ for } u \in [t_{i-1}, t_i), i = 1, \dots, m \quad (2.5.1)$$

where $0 = t_0 < t_1 < \dots < t_m = +\infty$.

Model-1

For $t \in [t_{i-1}, t_i)$ equation (2.2.1) becomes

$$f(t, x) = \sum_{j=1}^{i-1} r(x_j)(t_j - t_{j-1}) + r(x_i)(t - t_{i-1}) \quad (2.5.2)$$

and the resource is

$$R = \sum_{j=1}^{i-1} r(x_j)(t_j - t_{j-1}) + r(x_i)(T(x) - t_{i-1}),$$

if $T(x) \in [t_{i-1}, t_i)$.

Now define $T_k(x)$ as the life of an item tested under stress x in the interval $[t_{i-1}, t_i)$

$$T_k(x) = \begin{cases} 0 & \text{if } T(x) < t_{k-1} \\ T(x) - t_{k-1} & \text{if } t_{k-1} \leq T(x) < t_k \\ t_k - t_{k-1} & \text{if } T(x) \geq t_k \end{cases} \quad (2.5.3)$$

If model-1 holds on E then

$$\int_0^{t_p(x)} \frac{du}{t_p(x(u))} = 1 \quad (2.5.4)$$

and if the means $E(T(x)), E(T(x(u)))$ exists then

$$E\left(\int_0^{T(x)} \frac{du}{E(T(x(u)))}\right) = 1. \quad (2.5.5)$$

Here formula (2.5.5) determines the linear accumulation of damages (see Miner [60]). So in the case of stepwise stress of the form (2.5.1), formulas (2.5.4) and (2.5.5) imply the equality

$$\sum_{k=1}^m \frac{E(T_k(x))}{E(T(x_k))} = 1$$

and

$$\sum_{i=1}^{k-1} \frac{(t_i - t_{i-1})}{tp(x_i)} + \frac{tp(x) - t_{k-1}}{tp(x_k)} = 1, \quad (2.5.6)$$

if $tp(x) \in [t_{k-1}, t_k)$.

The inequality (2.5.6) is called Peshes-Stepanova model (see Kartashov [54]). These formulas can be used for estimation of the mean life under normal stress from the data of accelerated testing.

Model-2

Consider model-2 holds on E , then

$$R(x(t)) = G\left\{\int_0^t H(R_{x(u)}(u))d\log H(R_{x_0(u)})\right\},$$

where x is of the form in (2.5.1). Then for $t < t_1$ we have

$$H(R(x(t))) = H(R(x_1(t))).$$

Thus for $t \in [t_{k-1}, t_k)$ for $k > 1$

$$\begin{aligned} H(R(x(t))) &= \sum_{i=1}^{k-1} \frac{H(R(x_i(t)))}{H(R(x_1(t)))} [H(R(x_0(t_i))) - H(R(x_0(t_{i-1})))] \\ &+ \frac{H(R(x_k(t)))}{H(R(x_1(t)))} [H(R(x_0(t))) - H(R(x_0(t_{i-1})))]. \end{aligned} \quad (2.5.7)$$

This equality can be used for the estimation of the survival function under the normal stress from accelerated life experiments.

Model-3

Consider model-3 to be true on E , then

$$R(x(t)) = G\left\{\int_0^t H(R_{x(u)}(t))du\right\}$$

$$= G\{H(R(x_0(t))) + \int_0^t [HR_{x(u)}(u) - H(R(x_0(u)))]d \log u\}. \quad (2.5.8)$$

Now consider the stepwise stress of the form (2.5.1). Then for $t \in [t_{k-1}, t_k]$ we have

$$H[R(x(t))] = \frac{1}{t} \left\{ \sum_{i=0}^{k-1} H[R(x_i(t))](t_i - t_{i-1}) + H[R(x_k(t))](t - t_{k-1}) \right\}. \quad (2.5.9)$$

If the stress has the form

$$x(u) = \begin{cases} x_0 & \text{if } u \in [t_{2k}, t_{2k+1}) \\ x_1 & \text{if } u \in [t_{2k-1}, t_{2k}) \quad (k = 0, 1, \dots, m-1) \end{cases}$$

Here x_0 is the normal stress and x_1 is the accelerated stress , then

$$H(R(x(t))) = \frac{1}{t} \left\{ H(R(x_0(t))) \left[\sum_{i=0}^{k-1} (t_{2i+1} - t_{2i}) + t - t_{2k} \right] \right. \\ \left. + H(R(x_1(t))) \left\{ \sum_{i=0}^{k-1} (t_{2i+2} - t_{2i+1}) \right\} \right\}. \quad (2.5.10)$$

Therefore for $t \in [t_{2k+1}, t_{2k+2})$ equation (2.5.10) becomes

$$H(R(x(t))) = \frac{1}{t} \left\{ H(R(x_0(t))) \sum_{i=0}^{k-1} (t_{2i+1} - t_{2i}) + H(R(x_1(t))) \right\} \\ \cdot \left[\sum_{i=0}^{k-1} (t_{2i+2} - t_{2i+1}) + t - t_{2k+1} \right]. \quad (2.5.11)$$

when $m = 1$ then

$$H(R(x(t))) = \begin{cases} H(R(x(t))) & \text{if } t < t_1, \\ \frac{1}{t} [H(R(x_0(t)))t_1 + H(R(x_1(t)))(t - t_1)] & \text{otherwise.} \end{cases}$$

Model-5

Suppose model-5 is true on E then the transformational function satisfy the equality

$$f(t, x) = f(t - t_{i-1} + t_{i-1}^*, x_i), \quad (2.5.12)$$

if $t \in [t_{i-1}, t_i)$ for $i = 1, 2, \dots, m$

Here t_i^* can be found by solving the equations

$$\begin{aligned} f(t_1, x_1) &= f(t_1^*, x_2) \\ &\vdots \\ f(t_i - t_{i-1} + t_{i-1}^*, x_i) &= f(t_i^*, x_{i+1}) \end{aligned}$$

for $t_0 = 0$ and $i = 1, 2, \dots, m - 1$.

These equations are equivalent to

$$R(x(t)) = R(x_i(t - t_{i-1} + t_{i-1}^*)) \text{ if } t \in [t_{i-1}, t_i), \quad (2.5.13)$$

where $R(x_1(t)) = R(x_2(t_1^*))$.

These equations imply that

$$P[T(x) \geq t_1 + s | T(x) \geq t_1] = P[T(x_2) \geq t_1^* + s | T(x_2) \geq t_1^*]. \quad (2.5.14)$$

This equality is the so called model of Sedyakin (see Sedyakin [79]).

2.6 Accelerated Life Models when the Stress is Progressive

Suppose the stress x is progressive, *i.e.* it is continuously increasing in time t . For example, this stress is linear unidimensional of the form $x(t) = bt$. The models 1, 2 and 3 can be written as

$$R(x(t)) = G\left\{\frac{1}{b} \int_0^{bt} r(s) ds\right\} \quad (2.6.1)$$

$$R(x(t)) = G\left\{\int_0^t r(bu) dH(R_0(u))\right\} \quad (2.6.2)$$

$$R(x(t)) = G\left\{H(R_0(t)) + \int_0^t a(bu) du\right\} \quad (2.6.3)$$

For model-2, the parametrization of $r(bu)$ is $e^{\alpha+\beta bu}$. Now taking , $R(t) = G(e^\alpha H(R_0(t)))$ for

eliminating α , the model is

$$R(x(t)) = G\left\{\int_0^t e^{\beta b u} dH(R_0(u))\right\}. \quad (2.6.4)$$

Similarly, model-3 is

$$R(x(t)) = G\{H(R_0(t)) + \beta b t^2\}. \quad (2.6.5)$$

3 Possible Survival Distributions and Their Properties

3.1 Introduction

In this chapter, we consider some statistical distributions that are used to describe survival times in accelerated life testing. The distributions discussed in this thesis are the Gamma, Log-normal and Inverse Gaussian.

Consider a set of observations from a population. The natural question arises about the nature of the parent population. The clear idea is provided by the frequency polygon or frequency curve. But the information may be totally inadequate and unreliable, *i.e.* the sample observations may not cover the entire range of the parent distribution. Basically, an unusually high frequency in any one class may arise and the frequency curve may completely out of shape. In order to determine the frequency curve, we will have to make use of the technique of curve fitting to the given data. Some standard distributions may be tried and the best fitting distribution may be selected. Below, some important properties of common distributions used in survival and reliability analysis are described.

3.2 Gamma distribution

3.2.1 Gamma distribution as a model for failure times

In this thesis we consider the Gamma distribution as a model for failure times. We consider here the Gamma distribution because of its appealing features:

- (i) It accommodates a variety of shapes similar to the Weibull, Inverse Gaussian and log-normal distribution.
- (ii) This distribution has the structure of an exponential family and its many properties are associated with sampling distributions.
- (iii) Its derivation from a stochastic formulation of the failure process provides a physical support to its exact fit.

In this thesis, we consider a reciprocal linear regression model for the mean of the Gamma distribution. We consider the following situation. Suppose n objects are subjected to stress levels

x_1, x_2, \dots, x_n and their failure times are observed. Each object has a characteristic outset which we take to be the same. The stress levels signify the accumulation of the fatigue process at different levels. The Gamma distribution is a good competitor to other well known survival distributions such as the Inverse Gaussian, the Weibull and the Log normal.

3.2.2 Properties

The general form for the p.d.f. of a random variable x has a Gamma distribution

$$f_x(x) = \frac{1}{\beta^\alpha \Gamma \alpha} (x-r)^{\alpha-1} \exp\left[-\frac{(x-r)}{\beta}\right] \quad \text{for } x > r, \alpha > 0, \beta > 0, \quad (3.2.2.1)$$

where r and β are location and scale parameter respectively and α determines the shape of the p.d.f. This distribution is Pearson's type 3. The standard form of this distribution is obtained by putting $\beta = 1$ and $r = 0$ which is given by

$$f_x(x) = \frac{x^{\alpha-1} e^{-x}}{\Gamma \alpha} \quad \text{for } x \geq 0. \quad (3.2.2.2)$$

If $\alpha = 1$, the Gamma distribution reduces to the exponential distribution. If α is a positive integer, the Gamma distribution is sometimes called a special Erlangian distribution. Put $y = -x$ in (3.2.2.1) and (3.2.2.2) then,

$$f_y(y) = \frac{(-y-r)^{\alpha-1} \exp\left(\frac{y+r}{\beta}\right)}{\beta^\alpha \Gamma \alpha} \quad \text{for } y \leq -r. \quad (3.2.2.3)$$

$$f_y(y) = \frac{(-y)^{\alpha-1} e^y}{\Gamma \alpha} \quad \text{for } y \leq 0. \quad (3.2.2.4)$$

(3.2.2.3) and (3.2.2.4) are also Gamma distributions rarely considered and not discussed further here.

The cumulative distribution function of (3.2.2.2) is

$$P(X \leq x) = \frac{1}{\Gamma \alpha} \int_0^x e^{-t} t^{\alpha-1} dt, \quad (3.2.2.5)$$

which is called an incomplete Gamma function ratio but

$$\Gamma_x(\alpha) = \int_0^x e^{-t} t^{\alpha-1} dt, \quad (3.2.2.6)$$

is sometimes called incomplete Gamma function. Pearson [73] found it more convenient to use

$u = x\alpha^{\frac{1}{2}}$ instead of x and defined the incomplete Gamma function as

$$I(u, \alpha - 1) = \frac{1}{\Gamma(\alpha)} \int_0^{u\sqrt{\alpha}} t^{\alpha-1} e^{-t} dt.$$

The Gamma distribution with positive integer α can be derived as the distribution of the waiting time to the α th arrival from a Poisson source with parameter α . So it is apparent that the sum of k independent exponential variates with failure rate α has the Gamma distribution with parameters α and k . The continuous random variable x which is distributed according to the probability law

$$f(x) = \frac{e^{-x}x^{\alpha-1}}{\Gamma(\alpha)} \text{ for } \alpha > 0, 0 < x < \infty,$$

is known as a Gamma variate with parameter α and its distribution is called the Gamma distribution. The mean and variance of this distribution are equal to α , like in a Poisson distribution. The density function which can be seen to be a member of the exponential family is unimodal, positively skewed and Leptokurtic, with its mode at $x = \alpha - 1$ if $\alpha \geq 1$. But distribution (3.2.2.1) has a mode at $x = r + \beta(\alpha - 1)$. If $\alpha < 1$, $f_x(x)$ tends to infinity as x tends to zero, also if $\alpha = 1$, $\lim_{x \rightarrow 0} f_x(x) = 1$.

The m.g.f. can be easily found. Now the m.g.f. of the Gamma distribution (3.2.2.2) is

$$\begin{aligned} M_x(t) &= E(e^{tx}) = \int_0^{\infty} e^{tx} f(x) dx = \frac{1}{\Gamma\alpha} \int_0^{\infty} e^{tx} e^{-x} x^{\alpha-1} dx \\ &= (1-t)^{-\alpha}, |t| < 1. \end{aligned}$$

Thus the cumulant generating function $K_x(t)$ is given by

$$\begin{aligned} K_x(t) &= \ln M_x(t) = \ln(1-t)^{-\alpha} = -\alpha \log(1-t) \\ &= \alpha \left[t + \frac{t^2}{2} + \frac{t^3}{3} + \frac{t^4}{4} + \dots \right] \end{aligned}$$

From which we can derive

$$\text{Mean} = K_1 = \text{coefficient of } t \text{ in } K_x(t) = \alpha$$

$$\mu_2 = K_2 = \text{coefficient of } \frac{t^2}{2!} \text{ in } K_x(t) = \alpha$$

$$\mu_3 = K_3 = \text{coefficient of } \frac{t^3}{3!} \text{ in } K_x(t) = 2\alpha$$

$$K_4 = \text{coefficient of } \frac{t^4}{4!} \text{ in } K_x(t) = 6\alpha$$

$$\text{Therefore } \mu_4 = K_4 + 3\mu_2^2 = 6\alpha + 3\alpha^2$$

$$\text{Hence, } \beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{4\alpha^2}{\alpha^3} = \frac{4}{\alpha}$$

$$\text{and } \beta_1 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{6}{\alpha}$$

The moments can be found from either the m.g.f. or c.f. or directly by integration. From distribution (3.2.2.2) the r th moment about origin zero is

$$\mu_r' = (\Gamma\alpha)^{-1} \int_0^{\infty} x^{\alpha+r-1} e^{-x} dx = \frac{\Gamma(\alpha+r)}{\Gamma\alpha} \text{ for } r = 1, 2, \dots$$

The moments are found easily. Hence for distribution (3.2.2.2) has a Mean=variance= α . $\mu_3 = 2\alpha$ and $\mu_4 = 3\alpha^2 + 6\alpha$. The mean deviation of distribution (3.2.2.2) is $\frac{2\alpha^\alpha e^{-\alpha}}{\Gamma(\alpha)}$ and the coefficient of variance is $\frac{1}{\sqrt{\alpha}}$.

3.2.3 Applications and Uses

In 1900, Pearson [74] used an approximate chi-square statistics for various tests in contingency tables. But in this case, the exact distribution of this statistic is discrete. The Gamma distribution is approximately distributed as a positive definite quadratic forms and it is multinormally distributed variables. In 1938, the result of Welch [101] was proposed to Gamma distributions as a test criterion for the difference between expected values of two normal populations with possibly different variances. This distribution is used in place of the normal distribution as a parent distribution which is an expansion of Gram-Charlier series with Laguerre polynomial multipliers. Khamis [50] described this expansions and their properties. Barton [13] and Tiku [90], [91] used Laguerre series to approximate the distributions of smooth test statistics and non-central F .

The Gamma distribution gives us useful representation of many physical situations. It is used to make realistic adjustments to exponential distributions in representing life times in life testing situations. Now Weibull distributions have been more popular for this purpose although they do

not provide a permanent solution. Weibull families have simple forms of the failure rate function. Also the sum of independent exponentially distributed random variables represents a Gamma distribution which leads to the appearance in the theory of random counters and other related topics in associated with random process in meteorological precipitation process. This was discussed by Kotz and Neumann [49], Das [25].

3.3 Log-Normal Distribution

3.3.1 Introduction

Many distributions are far from being normal. However, in certain cases, logarithmic transformations of a variable make the distribution almost normal and thus changes a skewed distribution into a symmetrical distribution. The log-normal distribution is sometimes called the anti-log-normal distribution because it is not the distribution of the logarithm of a normal variable but of an exponential that is an antilogarithmic function of such a variable. When this distribution is applied to economic data, particularly production functions it is often called the Cobb-Douglas distribution (Dhrymes [24]).

In 1879, Galton [35] pointed out that if X_1, X_2, \dots, X_n are independent positive random variables and $T = \prod_{i=1}^n X_i$ then $\log T = \sum_{i=1}^n \log X_i$ would tend to a normal distribution as n tends to infinity .

For a non-negative random variable X define $U = \frac{\log(X-\theta) - \rho}{\sigma}$ where θ, ρ and σ are parameters and assume it has a standard normal distribution. Then the probability density function of X is

$$f_X(x) = [(x - \theta) \sqrt{2\pi} \sigma]^{-1} \exp\left[-\frac{1}{2} \left\{ \frac{\log(x - \theta) - \rho}{\sigma} \right\}^2\right], \quad (3.3.1.1)$$

which is known as three parameter Log-Normal distribution. The name 'Log-Normal' can also be applied to the distribution of X if $\log(\theta - X)$ is normally distributed, X having zero probability of exceeding θ .

3.3.2 Properties

The r th moment of X about zero is

$$\mu_r' = E(x^r) = \exp(r\rho + \frac{1}{2}r^2\sigma^2). \quad (3.3.2.1)$$

This moment sequence $\{\mu_r'\}$ is not unique to the lognormal distribution. Heyde [42] has shown that the distribution can not be defined by its moments. In 1917, Wicksell [94] obtained formulas

for the higher moments while Van Uven [92], [93] considered transformations to normality from a more general point of view. The mean of this distribution is

$$\mu_1 = \exp\left(\rho + \frac{1}{2}\sigma^2\right)$$

and when $\omega = e^{\sigma^2}$ the lower order central moments are

$$\begin{aligned}\mu_2 &= e^{2\rho} e^{\sigma^2} (e^{\sigma^2} - 1) = \omega(\omega - 1)e^{2\rho} \\ \mu_3 &= \omega^{\frac{3}{2}} (\omega - 1)^2 (\omega + 1) e^{3\rho} \\ \mu_4 &= \omega^2 (\omega - 1)^2 (\omega^4 + 2\omega^3 + 3\omega^2 - 3) e^{4\rho} .\end{aligned}$$

The coefficient of variation is $(\omega - 1)^{1/2}$. This distribution of X is unimodal and its mode is

$$\exp(\rho - \sigma^2) = \int_{\theta}^M P_X(x) dx = \frac{1}{2},$$

and median of X is e^ρ .

$$\therefore \text{Mean} > \text{Median} > \text{Mode}.$$

As $\sigma \rightarrow 0$ or δ to infinity the standard Log- Normal distribution tends to a unit normal distribution. Actually as σ increases, the log-normal distribution rapidly becomes markedly non-normal. Wise [95] has shown that the probability density function of the two parameter distribution has two points of inflection at $x = \exp\left[\rho - \frac{3\sigma^2}{2} \pm \sigma\sqrt{1 + \frac{1}{4}\sigma^2}\right]$.

3.3.3 Applications and Uses

The Log-normal distribution is usually represented as the distribution of various economical variables (see Gibrat [38], [39]). Gaddum [36] and Bliss [14] found that the distribution of the critical dose for a number of forms of drug application could be represented with adequate accuracy by a two parameter Log-normal distribution. In 1937-1940, Cochran [17], Williams [96] [97], Grundy [40], Herdan [44], [45] and Pearce [75] described the use of the Log-normal distribution in agricultural, entomological and literary research. Wu [98] has shown that Log-normal distributions can arise as limiting distributions of order statistics if sample size and order increase in certain relationships. Kolmogrov [53], Tomlinson [88], Oldham [72] applied this distribution of particle sizes in natural aggregates and in the closely related distribution of

dust concentration in industrial atmospheres. Geological applications have been described by Ahrens [2], Chayes [18], Miller and Goldberg [61] and Prohorov [76], Oldham [72] described the deviation of sickness absence and physicians consultation time. Wise [95] has described the application to dyedilution curves representing the concentration of the indicator as a function of time. This distribution gives a good representation of flood flows although extreme value distributions are more generally associated with this field (see Hermanson and Johnson [43]). Ferrell [29], Morrison [62] and Rohn [78] describe the use of this distribution in quality control. Koopmans et al. [55] pointed out that if the normal distribution is replaced by a log-normal distribution then confidence limits for the coefficients of variations are easily constructed. Wise [95] has pointed out marked similarities in shape between appropriately chosen log-normal distributions, random walks and Gamma distributions. The log-normal distribution is also applied in certain approximations to the distributions of Fisher's $Z = \frac{1}{2} \log F$. It is a well known approximation because the distribution of Z is much closer to normality than that of F (Aroian [1], Curtiss [21], Pearce [75]). Logarithmic transformations are often used in attempts to equalize variances.

3.4 Inverse Gaussian distribution

3.4.1 Introduction

The inverse Gaussian distribution was derived first by Schrodinger [85] as the probability distribution function of the first passage time in Brownian motion. Tweedie [87] proposed the name inverse Gaussian distribution since he found an inverse relationship between the cumulant generating functions of this distribution and those of Gaussian distributions. But in sequential analysis, Wald [99] derived an asymptotic form of the distribution of average sample numbers, which is the limiting form of the distribution of sample size in a sequential probability ratio tests. Because of this derivation, this distribution is sometimes called Wald's distribution.

The probability density function of a random variable X distributed as inverse Gaussian distribution with parameters μ and λ is denoted by $IG(\mu, \lambda)$ and given by

$$f(x, \mu, \lambda) = \left(\frac{\lambda}{2\pi x^3} \right)^{\frac{1}{2}} \exp(-\lambda(x - \mu)^2 / (2\mu^2 x)) \text{ for } x > 0. \quad (3.4.1.1)$$

The Inverse Gaussian distribution belongs to a two parameter family of distributions. The p.d.f.

of this distribution can be represented in several different forms each of which would be convenient for some purpose. Another important form can be obtained by a Wiener process w , in one dimension with positive drift ν and $w(0) = 0$. The time T which is required for the Wiener process to reach an arbitrary real value a , is a random variable with d.f.

$$f(t) = \frac{a}{\sigma\sqrt{2\pi t^3}} \exp\left\{-\frac{(a-\nu t)^2}{2\sigma^2 t}\right\} \text{ for } t > 0, \nu > 0. \quad (3.4.1.2)$$

This form is a reparametrization of the p.d.f. in (3.4.1.1) obtained by putting $\mu = \frac{a}{\nu}$ and $\lambda = \frac{a^2}{\sigma^2}$, where a is specified. There are various other representations of the density function but we will use the p.d.f. in (3.4.1.1).

3.4.2 Properties

The mean and variance of this distributions are μ and $\frac{\mu^3}{\lambda}$ respectively. This distribution is unimodal and positively skewed, with mode

$$x_{\text{mode}} = \mu \left\{ \left(1 + \frac{9}{4\phi^2} \right)^{\frac{1}{2}} - \frac{3}{2\phi} \right\}$$

and its shape depends only on the value of $\phi = \frac{\lambda}{\mu}$. The cumulant generating function of this distribution is obtained as

$$K_x(t, \mu, \lambda) = \frac{\lambda}{\mu} \left\{ 1 - \left(1 + \frac{2\mu^2 t}{\lambda} \right)^{\frac{1}{2}} \right\}. \quad (3.4.2.1)$$

Therefore the first four cumulants are

$$K_1 = \mu$$

$$K_2 = \frac{\mu^2}{\lambda}$$

$$K_3 = \frac{3\mu^5}{\lambda^2}$$

$$K_4 = \frac{15\mu^7}{\lambda^3}.$$

Generally, for $r \geq 2$, Tweedie [13] found the formula for the r th cumulant *i.e.* $K_r = 1.3.5.....(2r-3)\phi^{-(r-1)}$ where $\phi = \frac{\lambda}{\mu}$. The characteristic function of this distribution is

$$\Phi_x(t) = E[e^{itx}] = \exp\left\{ \Phi \left(1 - \left(1 - \frac{2i\mu^2 t}{\lambda} \right)^{\frac{1}{2}} \right) \right\}. \quad (3.4.2.2)$$

The central moments can be derived from cumulants, by direct integration or by using the characteristic function. From these raw moments can easily be obtained. For this distribution, all positive and negative moments exists. There is a remarkable relation between negative and positive moments, given by

$$E[X^{-r}] = \frac{E[X^{r+1}]}{\mu^{2r+1}}. \quad (3.4.2.3)$$

As $\phi = \frac{\lambda}{\mu} \rightarrow \infty$ with fixed μ , the standardized inverse Gaussian distribution ($\mu = 1$) tends to a unit normal distribution. If $Z = \frac{\lambda X}{\mu^2}$ then $Z \sim IG(\phi, \phi^2)$, where $\phi = \frac{\lambda}{\mu}$. This transformation gives a single parameter family of distributions. Wasan and Roy [100] found that as $\mu \rightarrow \infty$ and λ is fixed then $Y = \frac{1}{X}$ tends to the gamma distribution, *i.e.*

$$f_Y(y) = \lambda(2\pi)^{-\frac{1}{2}} y^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\lambda y\right) \text{ for } y > 0.$$

3.4.3 Applications and Uses

In the area of engineering reliability, inverse Gaussian distribution have been used to model of the movements of particles in a colloidal suspension under electric field. This distribution was long known in the literature of stochastic process and its potential in statistical applications in increasingly recognized in recent years. Wasan [102] also explained that this distribution have been used to the motion of particles influenced by Brownian movements. Tweedie [87] suggested that since \bar{X} is a sufficient statistic for μ , then the statistical independence between \bar{X} and $\frac{1}{\lambda}$ can be performed for nested classification in analysis of variance and this analogue developed for the analysis of variance with the values of chi-square and F tables. Although this distribution calls as the first passage time distribution of Brownian motion with positive drift suggests that a wide variety of shapes generated by the p.d.f. makes a good competitor to the Gamma, Weibull and log-normal models. Since it is a positively skewed distribution and this distribution has advantage over some other skewed distributions because of exact small sample theory is tractable and in some cases it is parallels that of the normal distribution. Inverse Gaussian distribution is also used in the area of natural and social sciences, *i.e.*, lengths of strikes (Lancaster [58]), hospital stays (Etan and Whitmore [28]), employee service times (Chhikara and Folks [19]), noise intensity (Marcus [63]) and tracer dilution curves (Wise [95]) *etc.* Analytically

inverse Gaussian distribution is sufficient to use for curve fitting but its scientific interest is limited. This distribution serves as a good model for accelerated life tests. Bannerjee and Bhattacharyya [10] applied this distribution in marketing research and Chhikara and Folks [20] consider applications of IG in life testing. Bhattacharyya and Fries [5] argue that IG is more appropriate than the Birbaum Saunders fatigue distribution and Chhikara and Gultman [22] gives sequential and Bayesian prediction limits. So the IG widely used tool in reliability theory. Gerstein and Mandelbred [41] showed that the IG model provides a good fit for the spontaneous activity of several neurons in the auditory of a cat and explained that by introducing a time varying drift for the Brownian motion they replicate the behaviors of one of the neurons subjected to periodic stimuli of various frequencies. Weiss [103] has given a review of the various types of random walk models for physical systems. Bachelier [12] used IG in stock prices. Recently Bannerjee and Bhattacharyya [10] used this IG model in renewal process and Whitmore [104] used this model to find labour turnover in marketing and labour research area.

4 Models of Failure Times in Accelerated Life Testing

4.1 The Reciprocal Linear Regression Model

In stochastic modeling of failure times, the fatigue life time distribution plays to a prominent role in the engineering literature. Bhattacharyya and Fries [4] motivated their Inverse Gaussian reciprocal linear model as follows. For a given stress, they assumed that the fatigue grows to a level where the component fails. Assuming, the accumulated fatigue to be governed by a Wiener process, the time to failure is distributed as an Inverse Gaussian distribution. This fact is due to the particular characterization of the Inverse Gaussian distribution as a first passage time distribution (see Johnson and Kotz [47]). For a stochastic relation of y to the intensity of stress x , they assumed that the severity of the stress levels (x) does not change the form of the life time distribution (y) but the stress levels (x) have an influence on the value of the parameters. In this case, the parameter μ is taken most important role to have a direct relation to x because it measures the mean fatigue growth per unit time.

Since, in accelerated life testing, it may be assumed that higher stress produces smaller mean failure time, they considered the following simple model

$$\{\mu(x)\}^{-1} = \alpha + \beta x \text{ for } \alpha > 0, \beta \geq 0, x > 0 \quad (4.1.1)$$

In a practical situation, we only consider $\alpha + \beta x > 0$ on a finite interval of x which corresponds to the range of stress x . But we assume that the origin is taken at the lower point of this interval, *i.e.* $\alpha \geq 0$.

Since, the fatigue life distribution may not follow a Wiener process, Inverse Gaussian distribution may not be always appropriate. Hence, we would like to consider some other general family of distributions in accelerated life testing. Due to proximity of the Gamma distribution to the Log-normal and Inverse Gaussian family, we are inclined to use the Gamma distribution as the model for failure times, *i.e.* the failure times at stress level x may be assumed to follow Gamma(α_x, β_x) distribution. The choice of α_x, β_x for a given x may be motivated from the following considerations. A constant β_x implies that the distribution shape may change with respect to x but not the scale, where as, a constant α_x implies that the distribution changes with x

according to scale changes. Moreover, since, $\mu_x = \alpha_x \beta_x$, the mean failure time at level x is considered to be a decreasing function of x , a general model for μ_x may be given by

$$\mu_x = g(\alpha_x, \beta_x),$$

where g is an increasing function, assuming that α_x is decreasing in x for fixed β_x and β_x is decreasing in x for fixed α_x . For simplicity, we consider the following choices

$$\text{FirstModel} \quad \alpha_x = (\alpha_0 + \alpha_1 x)^{-1}$$

$$\text{SecondModel} \quad \beta_x = (\beta_0 + \beta_1 x)^{-1}$$

$$\text{ThirdModel} \quad \alpha_x = [\alpha_0 + \alpha_1 \left(\frac{1}{x}\right)]$$

$$\text{FourthModel} \quad \beta_x = [\beta_0 + \beta_1 \left(\frac{1}{x}\right)]$$

The method of estimation involving these models is maximum likelihood, a general treatment of which is given in the following section.

4.2 Introduction of ML Method

Method of Maximum Likelihood is the most widely used method of estimation which was initially formulated by C. F. Gauss but as a general method of estimation was first introduced by Professor R. A. Fisher. Method of maximum likelihood has the following attractive features:

- i)* Generally it is a simple procedure, although the computational problems may not always be simple.
- ii)* The asymptotic properties of maximum likelihood estimators for under certain regularity conditions make their use desirable.
- iii)* Maximum likelihood estimation affords a rather general methods of estimation of parameters of survival distribution's even when observations are censored for example one can in most instances obtain the maximum likelihood estimators of the parameters of the survival distribution.

For some models, explicit solution of the maximum likelihood estimators may be possible. However for other models solutions can not be obtained explicitly. For this reason, there are two computational methods of finding maximum likelihood estimators such as Newton-Raphson

method and Scoring method. These are described in the next section.

4.3 Computational Aspects of ML Method

Let us consider a random sample of n observations X_1, \dots, X_n from a population with density function $f(x, \theta)$. The joint density function of sample values when regarded as a function of unknown parameter θ is called likelihood function and is denoted by

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \dots f(x_n, \theta) = \prod_{i=1}^n f(x_i, \theta).$$

The principle of maximum likelihood consists in finding an estimator of the parameter which maximizes L for variations in the parameter. Thus if there exists a function $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ of the sample values which maximizes L , then $\hat{\theta}$ is to be taken as estimator of θ . Thus $\hat{\theta}$ is the solution of the equation $\frac{\partial L}{\partial \theta} = 0$ subject to the condition that $\frac{\partial^2 L}{\partial \theta^2} < 0$ i.e. θ consists of a single parameter. Since $L > 0$ so is $\log L$ which shows that L and $\log L$ attain their extreme values at the same value of $\hat{\theta}$. But it is more convenient to work with $\log L$. So M.L. estimator $\hat{\theta}$ is generally obtained as the solution of the equation $\frac{\partial \log L}{\partial \theta} = 0$ subject to the condition that $\frac{\partial^2 \log L}{\partial \theta^2} < 0$. For multiparameter case, the MLE's $\hat{\theta}_1, \dots, \hat{\theta}_k$ of $\theta_1, \dots, \theta_k$ are obtained as the solution of the $k \times k$ system of equations, i.e.

$$\frac{\partial \log L(\theta_1, \dots, \theta_k)}{\partial \theta_i} \Big|_{\theta_i = \hat{\theta}_i} = 0, \quad i = 1, 2, \dots, k$$

(See Cramer [15]). The estimators $(\hat{\theta}_1, \dots, \hat{\theta}_k)$ are asymptotically normally distributed with mean $(\theta_1, \dots, \theta_k)$ and variance-covariance matrix $V_{\hat{\theta}}$ (see Rao[77]) where

$$V_{\hat{\theta}} = \begin{bmatrix} -E\left(\frac{\partial^2 \log L}{\partial \theta_1^2}\right) & \dots & -E\left(\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_k}\right) \\ \vdots & \ddots & \vdots \\ -E\left(\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_k}\right) & \dots & -E\left(\frac{\partial^2 \log L}{\partial \theta_k^2}\right) \end{bmatrix}$$

i.e.

$$\begin{bmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_k \end{bmatrix} \sim N \left[\begin{bmatrix} \theta_1 \\ \vdots \\ \theta_k \end{bmatrix}, v_{\hat{\theta}} \right]$$

In cases where it is not possible to obtain Maximum Likelihood Estimators for the parameters of the distribution then we use numerical techniques. There are two cases of the numerical techniques

- i) No constraints on values of the parameters are assumed.
- ii) Parameters are subject to some constraints.

In the second case, maximizing the logarithm of the likelihood functions, the constraint is typically as follows

The value of a parameter must lie in the interior of a particular region and must not lie on the boundary of that region. Numerical procedures that do not allow for constraints can be used as long as successive maximum likelihood estimators lie in the interior of the region.

Here we can describe maximizing techniques that do not consider constraints. These techniques can be put into direct and indirect classes. In the direct class, starting value is determined which thought to be a good approximation to the desired value. An example of this class is the method of steepest ascent or gradient method of Cauchy. In the indirect class, at first the derivatives of the logarithm of the likelihood function with respect to each parameter are obtained and then equated to zero. Next, the values of the parameters are easily obtained in terms of the observations that simultaneously satisfy these equations. Two examples of this class are Newton-Raphson method and the method of Scoring which we investigate below. Rao [77] also discussed this indirect procedure. Therefore some modern and sophisticated methods are available for solving non-linear equations of the type which confronts us the Newton-Raphson method and the Scoring method are very practical to implement and calculations are not difficult.

4.4 Asymptotic Properties of Maximum Likelihood Estimators

Maximum Likelihood estimators are the most important methods of estimation because of their

asymptotic properties. Generally, the asymptotic properties for maximum likelihood estimates applies to samples with many failures and models that satisfy certain regularity conditions.

Let X_1, \dots, X_n be a random sample of size n from a population with density function $f(x)$. Then the likelihood function of the sample values x_1, \dots, x_n , usually denoted by L is given by

$$L = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta) \\ = \prod_{i=1}^n f(x_i, \theta).$$

The following assumptions are made

- i) The first and second order derivatives exist and are continuous functions of θ in a range \mathfrak{R} , including the true value of the parameter for almost all x . For every θ in \mathfrak{R} , $\frac{\partial \log L}{\partial \theta} < F_1(x)$ and $\frac{\partial^2 \log L}{\partial \theta^2} < F_2(x)$ where $F_1(x)$ and $F_2(x)$ are integrable functions over $(-\infty, \infty)$.
- ii) $\frac{\partial^3 \log L}{\partial \theta^3}$ exists such that $\frac{\partial^3 \log L}{\partial \theta^3} < M(x)$, where $E[M(x)] < k$, a positive quantity.
- iii) For every θ in \mathfrak{R} , $E(-\frac{\partial^2 \log L}{\partial \theta^2})$ is finite and non-zero.
- iv) The range of integration is independent of θ . But if the range of integration depends on θ then $f(x, \theta)$ vanishes at the extremes depending on θ .

Under the above conditions, the asymptotic properties of the maximum likelihood estimators are

- i) MLE is consistent.
 - ii) It is asymptotically normally distributed.
- i.e.*

$$\hat{\theta} \xrightarrow{a} N[\theta, \{I(\theta)\}^{-1}] \text{ where, } I(\theta) = -E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right]$$

- iii) It is asymptotically efficient and achieves the Cramer-Rao lower bound for consistent estimators.

The second property greatly facilitated hypothesis testing and the construction of interval estimates and the third property is a particularly powerful result.

Under certain general conditions, maximum likelihood estimators possesses some important theorem which can be helpful to prove the asymptotic properties of maximum likelihood estimators.

Cramer – Rao Theorem : “With probability approaching unity as n tends to infinity, the

likelihood equation $\frac{\partial \log L}{\partial \theta} = 0$, has a solution which converges in probability to the true value θ_0 " (Dugue [27]).

Huzurbazar's Theorem: Any consistent solution of the likelihood equation provides a maximum of the likelihood with probability tending to unity as the sample size tends to infinity (Huzurbazar [46]).

Cramer's Theorem : A consistent solution of the likelihood equation is asymptotically normally distributed about the true value θ_0 . Thus $\hat{\theta}$ is asymptotically $N(\theta_0, \frac{1}{I(\theta_0)})$ (Cramer [15]). Here $I(\theta_0)$ is known as the information on θ supplied by the sample X_1, \dots, X_n .

4.5 The Newton-Raphson Method

The Newton-Raphson method is a widely used and often studied method for minimization. When the derivative of $g(\theta)$ is a simple expression which can be easily found, the real roots of the equation $g(\theta) = 0$ can be found rapidly by a process called the Newton-Raphson method after its discoverers. This method consists in finding an approximate value of the desired root graphically or otherwise and then finding a correction term which must be applied to the approximate value to get the exact value of the root. We illustrate this technique by solving first for a single θ and then presenting the case for $\theta_i, i = 1, 2, \dots, k$. Let

$$g(\theta) = \frac{\partial \log L(\theta)}{\partial \theta}, \quad (4.5.1)$$

where

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta).$$

The problem is then to find the value of θ say $\hat{\theta}$ such that

$$g(\hat{\theta}) = 0. \quad (4.5.2)$$

Thus $\hat{\theta}$ is the requisite maximum likelihood estimator of θ . If there is more than one value $\hat{\theta}$ such that $g(\hat{\theta}) = 0$, the choice of an initial value is very important. In most cases, the initial value obtained is in the neighborhood of the maximum likelihood estimator. When in doubt, consider several different initial values.

If $\hat{\theta}$ can not be obtained explicitly from solving equation $g(\hat{\theta}) = 0$, we may attempt a solution by means of the Newton-Raphson procedure. Suppose $\hat{\theta}_0$ be an initial value of $\hat{\theta}$ and then finding a correction term h so that the equation $g(\hat{\theta}) = 0$ becomes $g(\hat{\theta}_0 + h) = 0$. Now expanding $g(\hat{\theta}_0 + h)$ by Taylor's theorem we have

$$g(\hat{\theta}_0 + h) = g(\hat{\theta}_0) + h \cdot g'(\hat{\theta}_0) + \frac{h^2}{2!} \cdot g''(\hat{\theta}_0) + \dots = 0.$$

Supposing h is very small, we may neglect the terms containing h^2 and other higher powers. Then we have

$$g(\hat{\theta}_0) + h \cdot g'(\hat{\theta}_0) = 0.$$

which implies that $h = -\frac{g(\hat{\theta}_0)}{g'(\hat{\theta}_0)}$.

Then the improved value of the root is given by

$$\theta_0^{(1)} = \hat{\theta}_0 - \frac{g(\hat{\theta}_0)}{g'(\hat{\theta}_0)}.$$

Here, $\theta_0^{(1)}$ is called first approximation of the desired root. Similarly in the same way, the v th approximation of $\hat{\theta}$ is given by

$$\theta_0^{(v)} = \hat{\theta}_{v-1} - \frac{g(\hat{\theta}_{v-1})}{g'(\hat{\theta}_{v-1})}, \quad (4.5.3)$$

where

$$g'(\hat{\theta}_{v-1}) = \left. \frac{dg(\theta)}{d\theta} \right|_{\theta=\hat{\theta}_{v-1}}. \quad (4.5.4)$$

Thus, Newton-Raphson method consists of solving each iteration, the equation is

$$g(\hat{\theta}_v) + (\theta - \hat{\theta}_v)g'(\hat{\theta}_v) = 0.$$

For the next iterate $\hat{\theta}_{v+1}$, the solution takes the form of equation (4.4.2) by replacing $\hat{\theta}$ by $\hat{\theta}_{v+1}$, and $\hat{\theta}_{v-1}$ by $\hat{\theta}_v$. It is evident from Newton-Raphson formula that the larger the derivative $g'(\theta)$, the smaller is the correction term which must be applied to get the correct value of the root. This means that when the graph of $g(\theta)$ is nearly vertical where it crosses the x -axis, the correct value of the root can be found rapidly. But this method should not be used when the graph of $g(\theta)$ is

nearly horizontal where it crosses the x -axis.

Suppose $f(x; \theta_1, \dots, \theta_k)$ is a density function containing k -parameters $\theta_1, \dots, \theta_k$, $k \geq 2$. Furthermore, suppose the maximum likelihood estimators $\hat{\theta}_1, \dots, \hat{\theta}_k$ of $\theta_1, \dots, \theta_k$ are respectively found by differentiating the logarithm of the likelihood function with respect to $\theta_1, \dots, \theta_k$ and then equating to zero and then solving the resulting equations in terms of $\theta_1, \dots, \theta_k$. This leads to a system of k equations in k unknown parameters which can not be solved directly. We then extend Newton-Raphson method to k -dimensions.

Suppose $L(\theta_1, \dots, \theta_k)$ be the likelihood function of the k parameters distribution and let the maximum likelihood estimators of $\theta_1, \dots, \theta_k$ be found by solving simultaneously the vector equation

$$\underline{g}(\hat{\theta}_1, \dots, \hat{\theta}_k) = \underline{0}, \quad (4.5.5)$$

where

$$\underline{g}(\hat{\theta}_1, \dots, \hat{\theta}_k) = (g_1(\hat{\theta}_1, \dots, \hat{\theta}_k), \dots, g_k(\hat{\theta}_1, \dots, \hat{\theta}_k))$$

and

$$g_i(\hat{\theta}_1, \dots, \hat{\theta}_k) = \frac{\partial \log L(\theta_1, \dots, \theta_k)}{\partial \theta_i} \Big|_{\theta_i = \hat{\theta}_i}, \quad i, j = 1, \dots, k \quad (4.5.6)$$

Let us consider the initial estimates of $\hat{\theta}_1, \dots, \hat{\theta}_k$ are respectively $\hat{\theta}_{10}, \dots, \hat{\theta}_{k0}$. Then the v th iteration of $\theta_{1v}, \dots, \theta_{kv}$ of the solution $\hat{\theta}_1, \dots, \hat{\theta}_k$ is

$$\underline{\hat{\theta}}'_v = \underline{\hat{\theta}}'_{v-1} - g^{(v-1)} \|\nu_{ij}\|_{(v-1)}^{-1}, \quad (4.5.7)$$

where $\underline{\hat{\theta}}'_v = \hat{\theta}'_{1v}, \dots, \hat{\theta}'_{kv}$, $\underline{\hat{\theta}}'_{v-1} = (\hat{\theta}_{1,v-1}, \dots, \hat{\theta}_{k,v-1})$ and $g^{(v-1)} = (g_1^{(v-1)}, \dots, g_k^{(v-1)})$, $g_i^{(v-1)} = g_i(\hat{\theta}_{1,v-1}, \dots, \hat{\theta}_{k,v-1})$, $i = 1, 2, \dots, k$. Also $\|\nu_{ij}\|_{(v-1)}$ is the $k \times k$ matrix whose ij th element is

$$\nu_{ij} = \frac{\partial g_i(\theta_1, \dots, \theta_k)}{\partial \theta_j} \Big|_{(\theta_1, \dots, \theta_k) = (\hat{\theta}_{1,v-1}, \dots, \hat{\theta}_{k,v-1})}, \quad (4.5.8)$$

for $i = 1, \dots, k$, $j = 1, \dots, k$.

The important part of this method is to choose the initial estimates $\hat{\theta}_{10}, \dots, \hat{\theta}_{k0}$ because the

Newton-Raphson method will converge to a value that is not the maximum of likelihood function $[L(\theta_1, \dots, \theta_k)]$. In 1975, Gross and Clark noted that although this method does not always guarantee a maximum, it is a safeguard in that more than one set of initial values is considered and it means that any particular values in convergence can be uncovered.

4.6 Convergence of the Newton-Raphson Method

For considering single θ , the Newton-Raphson method formula for finding the roots are

$$\begin{aligned}\theta_0^{(1)} &= \hat{\theta}_0 - \frac{g(\hat{\theta}_0)}{g'(\hat{\theta}_0)} \\ &\vdots \\ \theta_0^{(v)} &= \hat{\theta}_{v-1} - \frac{g(\hat{\theta}_{v-1})}{g'(\hat{\theta}_{v-1})}\end{aligned}\tag{4.6.1}$$

which shows that the Newton-Raphson method is really an iteration method. Since the above equation can be written symbolically in the form

$$\theta_{v+1} = \phi(u_v).$$

then from the condition of convergency of iteration process, we can say that the Newton-Raphson method converges when $|\phi'(u_v)| < 1$. Hence from equation (4.5.1) we have

$$\begin{aligned}\phi'(u_v) &= \frac{d}{dx} \left(x - \frac{g(x)}{g'(x)} \right) \Big|_{x=\hat{\theta}_{v-1}} \\ &= \left[1 - \frac{d}{dx} \left(\frac{g(x)}{g'(x)} \right) \right] \Big|_{x=\hat{\theta}_{v-1}} \\ &= \left[1 - \frac{g'(x)g'(x) - g(x)g''(x)}{(g'(x))^2} \right] \Big|_{x=\hat{\theta}_{v-1}} \\ &= \left[\frac{g(x)g''(x)}{(g'(x))^2} \right] \Big|_{x=\hat{\theta}_{v-1}} \\ &= \left[\frac{g(\hat{\theta}_{v-1})g''(\hat{\theta}_{v-1})}{(g'(\hat{\theta}_{v-1}))^2} \right]\end{aligned}$$

Therefore the sufficient condition for convergence is

$$\left| \frac{g(\hat{\theta}_{v-1})g''(\hat{\theta}_{v-1})}{(g'(\hat{\theta}_{v-1}))^2} \right| < 1,$$

therefore $|g(\hat{\theta}_{v-1})g''(\hat{\theta}_{v-1})| < (g'(\hat{\theta}_{v-1}))^2$.

Graphical representation has also proved that the Newton-Raphson method converges. See the figure in next page. We observed that since $g'(\hat{\theta}_0) \neq 0$, the tangent line is not parallel to the axis. Where this line crosses the θ axis, we find our next approximation $\hat{\theta}^{(1)}$ and so on. Generally we stop our iteration procedure when $\log L(\theta)$ stops increasing appreciably. Actually if any value of two approximations are fairly close, we will make no further approximation and accept the last approximation as the required value of $\hat{\theta}$. The value of $\log L(\theta)$ should be calculated at each step because it permits us to monitor the stopping procedure.

4.7 The Scoring Method

The method of scoring is similar to the Newton-Raphson method for obtaining maximum likelihood estimates of parameters which was established by C.R. Rao in 1952. Sometimes maximum likelihood equations are complicated so that the solutions can not be obtained directly. For this reason, a great mechanism is introduced by adopting the method known as the scoring method when the requisite system of equations for solutions is non-linear. In this method, consider a trial solution and derive linear equations for small additive corrections. This process will continue until the corrections become negligible. The difference between Newton-Raphson method and scoring method is that the matrix of second derivatives used in the Newton-Raphson technique is replaced by the matrix of the expected values of second derivatives in the method of scoring. We discuss this technique by solving first for a single θ and then presenting the case for θ_i ; $i = 1, 2, \dots, k$.

Suppose $L(\theta)$ be the likelihood function of the parameter θ then $\frac{\partial \log L}{\partial \theta}$ is defined as the efficient score for θ . Let θ_0 be the trial value of $\hat{\theta}$ then expanding $\frac{\partial \log L}{\partial \theta}$ by Taylor's expansion, we have

$$\begin{aligned}\frac{\partial \log L}{\partial \theta} &\approx \frac{\partial \log L}{\partial \theta_0} + (\theta - \theta_0) \frac{\partial^2 \log L}{\partial \theta_0^2} + \dots \\ &\approx \frac{\partial \log L}{\partial \theta_0} - \partial \theta I(\theta_0) + \dots\end{aligned}$$

where, $\partial \theta = \theta - \theta_0$ and $I(\theta_0)$ is the Fisher information at the value $\theta = \theta_0$ and it is the expected value of $\left(-\frac{\partial^2 \log L}{\partial \theta_0^2}\right)$. In large samples, the difference between $[-I(\theta_0)]$ and $\frac{\partial^2 \log L}{\partial \theta_0^2}$ will be very small which is negligible so that the correction term $\partial \theta$ is obtained from the equation

$$\partial \theta \cdot I(\theta_0) = \frac{\partial \log L}{\partial \theta_0}, \quad \left[\because E\left(\frac{\partial \log L}{\partial \theta}\right) = 0 \right]$$

therefore

$$\partial \theta = I^{-1}(\theta_0) \frac{\partial \log L}{\partial \theta_0}.$$

Therefore the first approximation is $(\theta_0 + \partial \theta)$. The process will continue until the stable value of θ is obtained.

Now consider the case of the simultaneous estimation of several parameters. Suppose $L(\theta_1, \dots, \theta_k)$ be the likelihood function of the parameters $\theta_1, \dots, \theta_k$. Let the i -th coefficient be defined by

$$S_i = \frac{\partial \log L}{\partial \theta_i}; \quad i = 1, 2, \dots, k.$$

and the Fisher information matrix is

$$I_{ij} = E(S_i S_j).$$

Let $\theta_1^0, \dots, \theta_k^0$ be the trial values of $\theta_1, \dots, \theta_k$ respectively and their small additive corrections $\partial \theta_1, \dots, \partial \theta_k$ are given by the following equations

$$\begin{aligned}I_{11}^0 \partial \theta_1 + \dots + I_{1k}^0 \partial \theta_k &= S_1^0 \\ &\vdots \\ I_{k1}^0 \partial \theta_1 + \dots + I_{kk}^0 \partial \theta_k &= S_k^0\end{aligned}$$

This process will continue with corrected values each time until stable values of $\theta_1, \dots, \theta_k$ are obtained. The main problem of this method is computation and inversion of the information

matrix at each stage of approximation. But after some stage, the information matrix may be kept fixed and only to recalculate the scores. When final values are reached at the final stage, the information matrix may be computed at the estimated values for obtaining the variances and covariances of estimates. A discussion of the Newton-Raphson method and the method of scoring is also given by Kale.

4.8 Review of Accelerated Life Testing

This section presents brief review of the work done in accelerated life testing.

4.8.1 Nelson's Work

Nelson [67], [68], [69] presented statistical methods for planning and analyzing accelerated life tests with the Arrhenius model. He suggested that when all test units are run to failure, Arrhenius model should be of wide interest for estimating life which as a function of temperature because test units are not limited to just accelerating variable but they are applicable to many accelerated life tests with other accelerating variables. The assumptions of this model for the population of product life times are as follows

- (i) For any temperature, the life distribution is log-normal.
- (ii) The logarithmic standard deviation of the distribution is constant.
- (iii) The $\mu(x)$ of the logarithmic life is a linear function of the reciprocal $x = \frac{1}{T}$ of the absolute temperature T that is

$$\mu(x) = \alpha + \beta x,$$

where α and β are parameters characteristic of the material and the test method. The important fact about a log-normal life distribution is that the logarithms of the times to failure have a normal distribution. And the logarithm of the log arithmetic mean life $\mu(x)$ is the median life and is regarded as a nominal life. He used log-normal distribution to find the proportion P of units failing at an absolute temperature T with reciprocal $x = \frac{1}{T}$ is

$$\begin{aligned}
 P &= \Phi\left(\frac{\log(t) - \mu(x)}{\sigma}\right) \\
 &= \Phi\left(\frac{\log(t) - \alpha - \beta x}{\sigma}\right)
 \end{aligned}$$

where $\Phi(\bullet)$ is the standard normal cumulative distribution. At each test temperature, Nelson obtained estimates of the median and other percentiles of the life distribution and then provides an estimate of the Arrhenius relationship between the median life and temperature. He used both graphical and analytical method and recommended that a combination of graphical and analytical methods be used for analysis of accelerated life test data since both methods compliment each other. In 1972, Nelson presented another approach for analyzing accelerated life test data with the Inverse Power Law model. He suggested that the Inverse Power Law model is satisfactory for describing life as a function of the accelerating variable if the relationship for the situation can be transformed into a linear one. The assumptions of the Inverse Power Law model are

- (a) For any constant positive stress, the life distribution is Weibull.
- (b) The shape parameter m of the Weibull distribution is constant *i.e.* independent of the stress.
- (c) The scale parameter θ is an inverse power function of the stress V *i.e.*

$$\theta(V) = \frac{1}{kV^n}, \quad (4.8.1.1)$$

where k and n are positive parameters characteristic of the material and the test method. Equation (4.8.1.1) is called the Inverse Power Law since the inverse of the stress is raised to the n th power. Under these assumptions, the fraction $F(t; v)$ of units failing by time $t > 0$ is

$$F(t; v) = 1 - \exp[-(kv^n t)^m].$$

Here Nelson used graphical method for analyzing accelerated life test results with the Inverse Power Law model when all test units are run to failure.

4.8.2 Work of Bhattacharyya and Fries

Bhattacharyya and Fries [4] discussed accelerated life testing with reciprocal linear regression model. They explained that a material fails when its accumulated fatigue exceeds a critical amount $w > 0$ and assumed that the fatigue growth take place over time according to a Weiner

process with drift $\mu > 0$. The physical aspect leads to a regression structure for the reciprocal of the mean θ . Then the first passage time through w has the Inverse Gaussian distribution $IG(\theta, \lambda)$ with $\theta = w\mu^{-1}$ and $\lambda = w^2\delta^{-2}$ where δ^2 denote the diffusion constant of the process. Actually the IG model conforms to the structure of an exponential family and the methodology of optimum statistical inferences including test of hypotheses is well developed. Consider the time of failure y_1, y_2, \dots, y_n corresponding to the stress setting x_1, x_2, \dots, x_n . Then $y_i; i = 1, \dots, n$ is distributed as $IG(\theta_i, \lambda)$ with $\theta_i^{-1} = \alpha + \beta x_i$. They claimed that a linear form of the mean life is a simple choice for describing life as a function of the accelerating variable. They studied here to estimate the parameters by maximum likelihood method and follow the asymptotic properties of the estimators. But this asymptotic theory is based on number of replicates go to infinity at a fixed rate and their analysis is to include a lack of fit test.

4.8.3 Work of Singpurwalla

Singpurwalla Nozer D.[83] presented the inference procedure for analyzing accelerated life tests with Arrhenius re-action rate model. He pointed out that the scale parameter of an exponential distribution is reparametrized as a function of the stress according to the Arrhenius re-action rate model. He proposed for meaningful inferences for the problem of accelerated life testing when the location and the scale parameters of an exponential failure distribution was reparametrized according to the power rule model. He considered the following situation. Suppose the device under consideration is subjected to the constant application of a single stress V_i and its failure distribution under V_i follows the shifted exponential

$$f(t; \lambda_i, \gamma_i) = \begin{cases} \lambda_i \exp(t - \gamma_i)\lambda_i; & \lambda_i > 0, t \geq \gamma_i, \gamma_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

By the Arrhenius model for $a \leq V_i \leq b$ where a and b are known as constants, the scale parameter λ_i is reparametrized by

$$\lambda_i = \exp\left(A - \frac{B}{V_i}\right),$$

where A and B are unknown parameters which will be estimated. Again he suggested that the location parameter γ_i is reparametrized as a linear function of the stress V_i as

$$\gamma_i = \alpha - \beta V_i \quad \text{for } a \leq V_i \leq b$$

where α and β are unknown parameters. He showed that the stress levels does not change the form of the life time distribution but the stress levels have an influence on the parameters. The general procedure here is to obtain estimates of the parameters A, B, α and β from the results of the accelerated life test and then use these estimates to obtain estimates of the scale and the location parameter at use stress conditions. This type of parametrization are appropriate of the large sample theory for the ML estimators by using the shapes of the maximum relative likelihood functions. The main goal of this idea deals with inference about mean life at use stress conditions.

4.8.4 Work of Babu and Chaubey

Babu and Chaubey [6] presented reciprocal linear regression model for analyzing regression variables and the observations on the dependent variable following inverse Gaussian distribution. To formulate the inverse Gaussian regression model let $y_i; i = 1, \dots, n$ be the time of failure is distributed as $IG(\mu_x, \lambda)$ where $\mu_x = x_i' \beta$ and $\beta = (\beta_1, \dots, \beta_p)'$ is a vector of regression parameter and $x_i = (x_{i1}, \dots, x_{ip})'$ be the vector of stress setting. Then we can write $y_i^{-1} = x_i' \beta + \epsilon_i$ where $\lambda^{-1} y_i \epsilon_i^2$ are i.i.d. $\chi_{(1)}^2$ variables. To estimate β and λ they used pseudo maximum likelihood estimators given by Whitmore [105] and Bhattacharyya and Fries [3]

$$\hat{\beta} = (X' Y X)^{-1} X' I$$

$$\hat{\lambda} = \frac{(I' Y^{-1} I - I' X \hat{\beta})}{n}$$

They established consistency and derived the asymptotic distribution of the psudo maximum likelihood estimators under very general assumptions on the design points.

4.9 Estimation of the Parameters of the Reciprocal Linear Regression Model by Maximum Likelihood Method under Gamma Failure Times

First model

In order to obtain estimates of the parameters, let us consider the observations

$(x_i, y_i); i = 1, 2, \dots, n$ from n runs of an accelerated life test experiment where y_i denote the failure time corresponding to the stress setting x_i . Here the random variables y_1, y_2, \dots, y_n are independent and they are distributed as Gamma distribution, such that $y_i \sim \text{Gamma}(\alpha_{x_i}, \beta)$ where

$$\alpha_{x_i}^{-1} = \beta(\beta_0 + \beta_1 x_i). \quad (4.9.1)$$

This gives a reciprocal linear model since the mean of the i th observation is $\alpha_{x_i} \beta$. We may also write

$$\alpha_{x_i}^{-1} = \alpha_0 + \alpha_1 x_i \quad \text{where, } \alpha_0 = \beta \beta_0 \text{ and } \alpha_1 = \beta \beta_1.$$

Now the corresponding likelihood function is given by

$$\begin{aligned} L(\alpha_0, \alpha_1, \beta) &= \prod_{i=1}^n \frac{e^{-\frac{y_i}{\beta}} y_i^{\alpha_{x_i}-1}}{\beta^{\alpha_{x_i}} \Gamma(\alpha_{x_i})} \\ &= e^{-\sum_{i=1}^n \frac{y_i}{\beta}} \prod_{i=1}^n \frac{y_i^{\alpha_{x_i}-1}}{\beta^{\alpha_{x_i}} \Gamma(\alpha_{x_i})}. \end{aligned} \quad (4.9.2)$$

Hence the log-likelihood function becomes

$$\log L = -\sum_{i=1}^n \frac{y_i}{\beta} + \sum_{i=1}^n (\alpha_{x_i} - 1) \log y_i - \sum_{i=1}^n \log \Gamma \alpha_{x_i} - \sum_{i=1}^n \alpha_{x_i} \log \beta, \quad (4.9.3)$$

where, $\alpha_{x_i} = \frac{1}{\alpha_0 + \alpha_1 x_i}$. So the maximum likelihood estimators of α_0, α_1 and β are given by the solution of the following equations

$$\frac{\partial \log L}{\partial \alpha_0} = 0, \quad \frac{\partial \log L}{\partial \alpha_1} = 0, \quad \frac{\partial \log L}{\partial \beta} = 0 \quad (4.9.4)$$

Now $\frac{\partial \log L}{\partial \alpha_0} = 0$, i.e.

$$-\sum_{i=1}^n \alpha_{x_i}^2 \log y_i + \sum_{i=1}^n \left\{ -r - \alpha_{x_i}^{-1} + \alpha_{x_i} \left(\sum_{p=1}^{\infty} \frac{1}{p(p + \alpha_{x_i})} \right) \right\} \alpha_{x_i}^2 + \sum_{i=1}^n \alpha_{x_i}^2 \log \beta = 0 \quad (4.9.5)$$

Similarly, $\frac{\partial \log L}{\partial \alpha_1} = 0$, i.e.

$$-\sum_{i=1}^n x_i \alpha_{x_i}^2 \log y_i + \sum_{i=1}^n \left\{ -r - \alpha_{x_i}^{-1} + \alpha_{x_i} \left(\sum_{p=1}^{\infty} \frac{1}{p(p + \alpha_{x_i})} \right) \right\} x_i \alpha_{x_i}^2 + \sum_{i=1}^n x_i \alpha_{x_i}^2 \log \beta = 0 \quad (4.9.6)$$

And $\frac{\partial \log L}{\partial \beta} = 0$, i.e.

$$\sum_{i=1}^n \frac{y_i}{\beta^2} - \sum_{i=1}^n \frac{\alpha_{x_i}}{\beta} = 0. \quad (4.9.7)$$

To solve the equation (4.9.5), (4.9.6) and (4.9.7) we will get the estimated value of α_0, α_1 , and β . But these equations are usually complicated to solve so that their solutions can not be obtained directly but their solutions can be numerically obtained by using the Newton-Raphson method. For various sets of data which we generated on a computer by some computer program it was found that $\hat{\alpha}_0, \hat{\alpha}_1$ and $\hat{\beta}$ could be obtained in a few iterations of the method.

Second model

Now we consider α is fixed that is y_1, \dots, y_n are independently distributed as Gamma (α, β_i) with

$$(\alpha \beta_i)^{-1} = \alpha_0 + \alpha_1 x_i, \quad (4.9.8)$$

$$\begin{aligned} \text{i.e. } \beta_i^{-1} &= \alpha(\alpha_0 + \alpha_1 x_i) \\ &= \beta_0 + \beta_1 x_i \end{aligned}$$

where $\beta_0 = \alpha \alpha_0, \beta_1 = \alpha \alpha_1$. So

$$\beta_i = \frac{1}{\beta_0 + \beta_1 x_i}.$$

In this case, the log-likelihood function is

$$\begin{aligned} L(\alpha, \beta_0, \beta_1) &= \prod_{i=1}^n \frac{e^{-\frac{y_i}{\beta_i}} y_i^{\alpha-1}}{\beta_i^{\alpha} \Gamma \alpha} \\ &= e^{-\sum_{i=1}^n \frac{y_i}{\beta_i}} \prod_{i=1}^n \frac{y_i^{\alpha-1}}{\beta_i^{\alpha} \Gamma(\alpha)} \end{aligned} \quad (4.9.9)$$

where $\beta_i = \frac{1}{\beta_0 + \beta_1 x_i}$. So

$$\log L = - \sum_{i=1}^n (\beta_0 + \beta_1 x_i) y_i + (\alpha - 1) \sum_{i=1}^n \log y_i - \alpha \sum_{i=1}^n \log \left(\frac{1}{\beta_0 + \beta_1 x_i} \right) - n \log \Gamma \alpha \quad (4.9.10)$$

Therefore the maximum likelihood estimators of β_0, β_1 and α are given by the solution of the following equations

$$\frac{\partial \log L}{\partial \beta_0} = 0, \quad \frac{\partial \log L}{\partial \beta_1} = 0, \quad \frac{\partial \log L}{\partial \alpha} = 0 \quad (4.9.11)$$

But these equations are complicated to solve so that their solutions can be numerically obtained by using the Newton-Raphson method. For several types of data which we generated on a computer program it was found that $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\alpha}$ could be obtained in a few iterations. The maximum likelihood estimators in third and fourth model can be similarly obtained.

5 Numerical Illustration

5.1 Data

Nelson [65] reports data on the failure of 40 motorettes with a new class-H insulation material in a motorette test performed at four elevated temperature setting run at $190^{\circ}C$, $220^{\circ}C$, $240^{\circ}C$ and $260^{\circ}C$. For each test temperature, the 10 motorettes were periodically examined for insulation failure and the given failure time is midway between the inspection time when the failure was found and the time of the previous inspection.

Table: Hours to failure for class-H insulation material

<u>190^oC</u>	<u>220^oC</u>	<u>240^oC</u>	<u>260^oC</u>
7228	1764	1175	600
7228	2436	1175	744
7228	2436	1521	744
8448	2436	1569	744
9167	2436	1617	912
9167	2436	1665	1128
9167	3108	1665	1320
9167	3108	1713	1464
10511	3108	1761	1608
10511	3108	1953	1896

Source: Nelson [65]

5.2 Estimates of the Parameters

For illustrative purposes, we fit the Gamma regression model to these data. Nelson [65] used the same data in the Arrhenius model by employing a combination of graphical and analytic techniques based upon the assumptions that the log-failure times are normally distributed with constant variance and the mean depends on temperature. Bhattacharyya and Fries [4] fit an Inverse Gaussian reciprocal linear model to be adequate for this data . They choose the x values given by $x = 10^{-8}(t^3 - 180^3)$, t denoting the temperature in centigrade. Babu and Chaubey [6] used this data to fit an Inverse Gaussian model given by $IG(\mu_x, \lambda)$ where $\mu_x^{-1} = \beta_0 + \beta_1 x$.

Actually the main purpose of the experiment was to estimate insulation life at $180^{\circ}C$ exceeded a minimum requirement. We took the distribution of failure times as Gamma *i.e.* $G(\alpha, \beta)$ with reciprocal linear regression model where $x = 10^{-8}(t^3 - 180^3)$. The computer program developed to estimate the parameters attached to the appendix was used for the purpose. We used here a SAS program PROC NLP [84] and this program provides estimates of the parameters, Hessian matrix, Covariance matrix, Correlation matrix and confidence limits for parameters and functions of them.

First step

Maximum likelihood estimates of the parameters

Model	Parameters	Estimates of the parameters	Standard error of the estimates
First	α_0	0.00433	0.00123
	α_1	0.77118	0.17239
	β	107.42389	24.11420
Second	β_0	0.00074293	0.00025522
	β_1	0.13956	0.03187
	α	19.18338	4.25277
Third	α_0	4.16547	1.17190
	α_1	0.69342	0.159
	β	124.755	28.04153
Fourth	β_0	23.07059	8.69414
	β_1	5.43162	1.28748
	α	17.19921	3.80919

Second step

CONFIDENCE INTERVAL FOR DIFFERENT PARAMETERS

An approximation to $100(1 - \alpha)\%$ confidence interval for α_i (say) as,

$$\hat{\alpha}_i \pm Z_{\frac{\alpha}{2}} [S.E.(\hat{\alpha}_i)]$$

where $Z_{\frac{\alpha}{2}}$ is the $100(1 - \alpha)$ -th percentile of the standard normal distribution.

An approximate 90% confidence interval for the parameters are shown in the next table.

Model	Parameters	Lower limit	Upper limit
First	α_0	0.00230665	0.00635335
	α_1	0.48759845	1.05476
	β	67.756031	147.091749
Second	β_0	0.0003230931	0.0011627669
	β_1	0.08713385	0.19198615
	α	12.18757335	26.17918665
Third	α_0	2.2376945	6.0932455
	α_1	0.431865	0.954975
	β	78.62668315	170.8833169
Fourth	β_0	8.7687297	37.3724503
	β_1	3.3137154	7.5495246
	α	10.93309245	23.46532755

5.3 Goodness of Fit of the Models

Since $Y_i \sim \text{Gamma}(\alpha_i, \beta)$ then

$$\widehat{Y}_i = \widehat{\alpha}_i \widehat{\beta} \quad (5.3.1)$$

Goodness of fit is judged using the following norms;

$$L_1 = \sum_{i=1}^n |\widehat{Y}_i - Y_i| \quad (5.3.2)$$

$$L_2 = \sum_{i=1}^n |\widehat{Y}_i - Y_i|^2 \quad (5.3.3)$$

For comparative purpose, we can use Bhattacharyya and Fries result and Babu and Chaubey's result. Bhattacharyya and Fries [4] find the Inverse Gaussian reciprocal linear model to be adequate for this data excluding the 260°C setting and they choose the x values given by, $x = 10^{-8}(t^3 - 180^3)$, t denoting the temperature in Centigrade. Babu and Chaubey used this data for both batches and they fit an Inverse Gaussian reciprocal linear model given by $IG(\mu_x, \lambda)$ where

$$\mu_x^{-1} = \beta_0 + \beta_1 x$$

They obtained to estimate the parameters by maximum likelihood method and are respectively given by

$$\widehat{\beta}_0 = 0.03731, \widehat{\beta}_1 = 7.317285, \widehat{\lambda} = 0.040233$$

Since $Y_i \sim IG(\mu_x, \lambda)$ then

$$\widehat{Y}_i = \frac{1}{\widehat{\mu}_{x_i}}$$

The models can be selected according the value of the likelihood and L_1, L_2 norms. In practice, L_1, L_2 norms are used but not be always appropriate. For this reason, we may also consider the likelihood as the selection criterion.

The table below shows the L_1 , L_2 norms and log likelihood function in the different model

Model	L_1	L_2	Log likelihood
First	19068.23003	18101706.78	306.1948749
Second	18937.04473	18119160.63	310.3946866
Third	19651.03725	19547356.61	309.2575996
Fourth	20946.02746	24074289.70	312.6183789
Babu and Chaubey	18784.80094	18231932.08	-

5.4 Conclusions and further Research

For the Gamma family, it can be observed that the reciprocal regression model for scale parametrization gives the better result than other types of model according to L_1 norm, but first model, *i.e.* shape parametrization is better according to L_2 norm. Thus we consider the likelihood criterion according to which the first model is the best. Note also that the estimators given in Babu and Chaubey [6] provide best model amongst all the models considered according to the L_1 norm this property is lost when we consider L_2 norm.

For further research, we can consider the use of the family of power transformations of dependent variable before carrying out the regression analysis. Below we give some more details of this.

Let Y_1, \dots, Y_n be a random sample from some non-normal density function. Let us assume that they follow a gamma distribution. Consider the following transformation

$$\underline{Y} \rightarrow \underline{Y}^{(\lambda)}$$

where $\underline{Y} = (Y_1, \dots, Y_n)'$ and $\underline{Y}^{(\lambda)} = (Y_1^{(\lambda)}, \dots, Y_n^{(\lambda)})'$. Here $Y_i^{(\lambda)}$ is the transformed value of Y_i for some $\underline{\lambda} = (\lambda_1, \dots, \lambda_k)$ from the family $\Lambda = (\lambda_1, \dots, \lambda_k)$ of transformation. From the above transformation we assume that $Y^{(\lambda)} \sim \text{Gamma}(\alpha_\lambda, \beta_\lambda)$. We will use the method of maximum

likelihood estimation for estimating λ from a given family. In 1964, Box and Cox proposed the parametric transformation which is

$$Y_i^{(\lambda)} = \begin{cases} \frac{Y_i^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log Y_i & \text{if } \lambda = 0 \end{cases} \quad (5.4.1)$$

But Tukey also suggested another type of transformation which is

$$Y_i^{(\lambda)} = \begin{cases} \frac{(Y_i + \lambda_2)^{\lambda_1} - 1}{\lambda_1} & \text{if } \lambda_1 \neq 0 \\ \log(Y_i + \lambda_2) & \text{if } \lambda_1 = 0 \end{cases} \quad (5.4.2)$$

For $\lambda = 1$, Box and Cox transformation follows a single shift in Y_i which affects the location but shape will be unchanged in the distribution. If λ varies over $(0,1)$ then it covers the range from $\log Y_i$ to $Y_i - 1$ or Y_i itself with something new in between. But if λ is not set in advance then for estimation we must of course prepare for values outside this range. This will investigate in future.

Appendix

The SAS System for First model

PROC NLP: Nonlinear Maximization

```
proc nlp data=final tech=tr vardef=n covariance=h pcov phes;
profile alpha0 alpha1 beta / alpha=.5 .1 .05 .01;
max loglik;
parms alpha0=.04, alpha1=7.4, beta=2892.4;
bounds beta >1e-12;
loglik=-(y/beta)+((1/(alpha0+(alpha1*x)))-1)*log(y)
-lgamma(1/(alpha0+(alpha1*x)))
-(1/(alpha0+(alpha1*x)))*log(beta);
run;
```

Optimization Start
Parameter Estimates

Parameter	Estimate	Gradient	Lower BC	Upper BC
1 ALPHA0	0.040000	816.18854	.	.
2 ALPHA1	7.400000	12.40821	.	.
3 BETA	2892.400000	-0.03093	1E-12	.

Value of Objective Function = -389.2812123

Hessian Matrix

	ALPHA0	ALPHA1	BETA
ALPHA0	-20134.66628	-236.0086413	0.291805417
ALPHA1	-236.0086413	-4.354286129	0.004882963
BETA	0.291805417	0.004882963	4.8600017E-6

Determinant = 0.333665073
Active Constraints= 0 Criterion= -306.19487
Maximum Gradient Element= 0.000106442 Lambda= 0 Rho= 0 Radius= 0.6205

Optimization Results
Parameter Estimates

	Approx. Parameter	t Estimate	Approx. Std Err	Ratio	Prob> t	Gradient Active BC
1	ALPHA0	0.00433	0.00123	3.54	0.0010	0.000106
2	ALPHA1	0.77118	0.17239	4.47	0.0001	7.0096E-7
3	BETA	107.42389	24.11420	4.45	0.0001	-5.109E-9

Value of Objective Function = -306.1948749

Hessian Matrix

	ALPHA0	ALPHA1	BETA
ALPHA0	-5668291.165	-67323.54382	705.80305958
ALPHA1	-67323.54382	-1308.268456	11.896872251
BETA	705.80305958	11.896872251	-0.113879592

Determinant = 4958235.1736

Covariance Matrix 2: H = (NOBS/d) inv(G)

	ALPHA0	ALPHA1	BETA
ALPHA0	1.5024515E-6	0.0001472442	0.0246943265
ALPHA1	0.0001472442	0.0297171712	4.0171093065
BETA	0.0246943265	4.0171093065	581.49459929

Factor sigm = 1

Determinant = 2.0168467E-7

Approximate Correlation Matrix of Parameter Estimates

	ALPHA0	ALPHA1	BETA
ALPHA0	1	0.69684	0.83546
ALPHA1	0.69684	1	0.96636
BETA	0.83546	0.96636	1

Determinant = 0.0077681796

Confidence Limits

Parameter	Estimate	Profile Likelihood			Wald	
		Alpha	Lower	Upper	Lower	Upper
ALPHA0	0.00433	0.500	0.00361	0.00526	0.00351	0.00516
	. 0.100	0.00279	0.00711	0.00232	0.00635	
	. 0.05	0.00257	0.00788	0.00193	0.00674	
	. 0.01	0.00221	0.00971	0.00118	0.00749	
ALPHA1	0.77118	0.500	0.66710	0.89910	0.65491	0.88745
	. 0.100	0.54579	1.14070	0.48763	1.05473	
	. 0.05	0.51294	1.23674	0.43331	1.10905	
	. 0.01	0.45667	1.45726	0.32714	1.215	
BETA	107.42389	0.500	92.87478	125.33121	91.15911	123.68867
	. 0.100	75.93396	159.22813	67.75957	147.08822	
	. 0.05	71.35430	172.73327	60.16093	154.68686	
	. 0.01	63.50486	203.80396	45.30983	169.53795	

The SAS System for Second Model

PROC NLP: Nonlinear Maximization

```
proc nlp data=final tech=tr vardef=n covariance=h pcov phes;
profile alpha0 alpha1 beta / alpha=.5 .1 .05 .01;
max loglik;
parms alpha0=.04, alpha1=7.4, beta=2892.4;
bounds beta >1e-12;
loglik=-(y/beta)+(((alpha0+(alpha1*(1/x))))-1)*log(y)
-lgamma(alpha0+(alpha1*(1/x)))
-((alpha0+(alpha1*(1/x)))*log(beta));
run;
```

Optimization Start
Parameter Estimates

Parameter	Estimate	Gradient	Lower BC	Upper BC
1 ALPHA0	0.040000	-209.12616	.	.
2 ALPHA1	7.400000	-7483	.	.
3 BETA	2892.400000	-3.54405	1E-12	.

Value of Objective Function = -45399.52324

Hessian Matrix

	ALPHA0	ALPHA1	BETA
ALPHA0	-0.347567308	-5.426975682	-0.013829346
ALPHA1	-5.426975682	-188.3970515	-0.481131939
BETA	-0.013829346	-0.481131939	0.001219465

Determinant = 0.0882049632

Optimization Results
Parameter Estimates

Parameter	Estimate	Approx. Std Err	t Ratio	Approx. Prob> t	Gradient	Active BC
1 ALPHA0	4.16547	1.17190	3.55	0.0010	7.8793E-9	
2 ALPHA1	0.69342	0.15900	4.36	0.0001	3.239E-7	
3 BETA	124.75500	28.04153	4.45	0.0001	1.3774E-9	

Value of Objective Function = -309.2575996

Hessian Matrix

	ALPHA0	ALPHA1	BETA
ALPHA0	-2.547743794	-44.19306969	-0.320628431
ALPHA1	-44.19306969	-1773.006194	-11.15487168
BETA	-0.320628431	-11.15487168	-0.072706742

Determinant = 3.2609032319

Covariance Matrix 2: H = (NOBS/d) inv(G)

	ALPHA0	ALPHA1	BETA
ALPHA0	1.3733438358	0.1114522181	-23.15560055
ALPHA1	0.1114522181	0.0252799772	-4.370016367
BETA	-23.15560055	-4.370016367	786.32757125

Factor sigm = 1

Determinant = 0.3066635005

Approximate Correlation Matrix of Parameter Estimates

ALPHA0 ALPHA1 BETA

ALPHA0	1	0.59815	-0.7046
ALPHA1	0.59815	1	-0.9802
BETA	-0.7046	-0.9802	1

Determinant = 0.0112331776

Confidence Limits

Parameter	Estimate	Alpha	Profile Likelihood		Wald	
			Lower	Upper	Lower	Upper
1 ALPHA0	4.16547	0.500	3.42386	4.98849	3.37503	4.95590
		. 0.100	2.44548	6.31774	2.23787	6.09307
		. 0.05	2.15802	6.78482	1.86859	6.46234
		- 0.01	1.63231	7.75029	1.14686	7.18408
2 ALPHA1	0.69342	0.500	0.59346	0.80547	0.58617	0.80066
		. 0.100	0.46306	0.98776	0.43189	0.95494
		. 0.05	0.42582	1.05211	0.38179	1.00504
		. 0.01	0.35843	1.18506	0.28387	1.10296
3 BETA	124.75500	0.500	107.83589		105.84127	143.66873
		. 0.100	88.14367	185.05844	78.63078	170.87922
		. 0.05	82.82266		69.79461	179.71539
		. 0.01	73.70203		52.52480	196.98520

The SAS System for Third Model
PROC NLP: Nonlinear Maximization

```
proc nlp data=final tech=tr vardef=n covariance=h pcov phes;
profile alpha0 alpha1 beta / alpha=.5 .1 .05 .01;
max loglik;
parms beta0=.04, beta1=7.4, alpha=2892.4;
bounds alpha > 1e-12;
loglik=-(y/(1/(beta0+(beta1*x))))+(alpha-1)*log(y)
-lgamma(alpha)-(alpha*log(1/(beta0+(beta1*x))));
run;
```

Optimization Start
Parameter Estimates

Parameter	Estimate	Gradient	Lower BC	Upper BC
1 BETA0	0.040000	258773	.	.
2 BETA1	7.400000	8726	.	.
3 ALPHA	2892.400000	-42.76325	1E-12	.

Value of Objective Function = -48974.51621

Hessian Matrix

	BETA0	BETA1	ALPHA
BETA0	-2441237.629	-40850.75962	138.27448705
BETA1	-40850.75962	-1599.824144	4.6579757457
ALPHA	138.27448705	4.6579757457	-0.013827162

Determinant = 4885.6211925

Optimization Results
Parameter Estimates

Parameter	Estimate	Approx. Std Err	t	Approx. Prob> t	Gradient	Active BC
1 BETA0	0.00074293	0.00025522	2.91	0.0059	1.0283E-7	
2 BETA1	0.13956	0.03187	4.38	0.0001	8.5502E-9	
3 ALPHA	19.18338	4.25277	4.51	0.0001	2.4555E-8	

Value of Objective Function = -310.3946866

Hessian Matrix

	BETA0	BETA1	ALPHA
BETA0	-45957951.41	-766891.8898	7359.0804364
BETA1	-766891.8898	-29929.03424	247.43744877
ALPHA	7359.0804364	247.43744877	-2.140430055

Determinant = 43533771758

Covariance Matrix 2: H = (NOBS/d) inv(G)

	BETA0	BETA1	ALPHA
BETA0	6.5138242E-8	4.1217113E-6	0.0007004309
BETA1	4.1217113E-6	0.0010156188	0.1315782877
ALPHA	0.0007004309	0.1315782877	18.086049033

Factor sign = 1

Determinant = 2.297067E-11

Approximate Correlation Matrix of Parameter Estimates

	BETA0	BETA1	ALPHA
BETA0	1	0.50675	0.64532
BETA1	0.50675	1	0.97084
ALPHA	0.64532	0.97084	1

Determinant = 0.019198321

Confidence Limits

Parameter	Estimate	Alpha	Profile Likelihood		Wald	
			Lower	Upper	Lower	Upper
1 BETA0	0.0007429	0.500	0.0005816	0.0009223	0.0005708	0.0009151
		. 0.100	0.0003691	0.00121	0.0003231	0.00116
		. 0.05	0.0003068	0.00131	0.0002427	0.00124
		. 0.01	0.0001929	0.00153	0.0000855	0.00140
2 BETA1	0.13956	0.500	0.11953	0.16202	0.11807	0.16106
		. 0.100	0.09339	0.19855	0.08714	0.19198
		. 0.05	0.08593	0.21144	0.07710	0.20202
		. 0.01	0.07242	0.23809	0.05747	0.22165
3 ALPHA	19.1838	0.500		16.31493	22.05182	
		. 0.100		27.05488	12.18819	26.17856
		. 0.05	12.02611	28.77568	10.84810	27.51865
		. 0.01		32.33112	8.22897	30.13778

The SAS System for Fourth Model

PROC NLP : Nonlinear Maximization

```
proc nlp data=final tech=tr vardef=n covariance=h pcov phes;
profile beta0 beta1 alpha / alpha=.5 .1 .05 .01;
max loglik;
parms beta0=.04, beta1=7.4, alpha=2892.4;
bounds alpha>1e-12;
loglik=-(y/((beta0+(beta1*(1/x)))))+(alpha-1)*log(y)
-lgamma(alpha)-(alpha*log((beta0+(beta1*(1/x)))));
```

run;

**Optimization Start
Parameter Estimates**

Parameter	Estimate	Gradient	Lower BC	Upper BC
1 BETA0	0.040000	-993.37999	.	.
2 BETA1	7.400000	-15543	.	.
3 ALPHA	2892.400000	-209.29233	1E-12	.

Value of Objective Function = -490512.0855

Hessian Matrix

	BETA0	BETA1	ALPHA
BETA0	11.782818164	133.37469585	-0.345497136
BETA1	133.37469585	2087.9220323	-5.403537853
ALPHA	-0.345497136	-5.403537853	-0.013827162

Determinant = -189.474342

Parameter Estimates

Parameter	Estimate	Approx. Std Err	t Ratio	Approx. Prob> t	Gradient	Active BC
1 BETA0	23.07059	8.69414	2.65	0.0114	8.8349E-7	
2 BETA1	5.43162	1.28748	4.22	0.0001	0.0000151	
3 ALPHA	17.19921	3.80919	4.52	0.0001	4.8547E-6	

Value of Objective Function = -312.6183789

Hessian Matrix

	BETA0	BETA1	ALPHA
BETA0	-0.064583156	-0.820764914	-0.34583265
BETA1	-0.820764914	-15.18155547	-5.895378867
ALPHA	-0.34583265	-5.895378867	-2.394605403

Determinant = 0.0211454195

Covariance Matrix 2: $H = (NOBS/d) \text{inv}(G)$

	BETA0	BETA1	ALPHA
BETA0	75.588132197	3.4715036805	-19.46319568
BETA1	3.4715036805	1.657614408	-4.582309994
ALPHA	-19.46319568	-4.582309994	14.509890445

Determinant = 47.291565895

Approximate Correlation Matrix of Parameter Estimates

	BETA0	BETA1	ALPHA
BETA0	1	0.31013	-0.5877
BETA1	0.31013	1	-0.9344
ALPHA	-0.5877	-0.9344	1

Determinant = 0.0260125204

Confidence Limits

Parameter	Estimate	Alpha	Profile Likelihood		Wald		
			Lower	Upper	Lower	Upper	
1 BETA0	23.07059	0.500	17.95599	29.59865	17.20648	28.93470	
			. 0.100	11.32680	42.51886	8.77000	37.37119
			. 0.05	9.26985	47.90117	6.03038	40.11080
			. 0.01	4.75244	60.77066	0.67596	45.46522
2 BETA1	5.43162	0.500	4.66056	6.39553	4.56322	6.30001	
			. 0.100	3.77287	8.26100	3.31389	7.54934
			. 0.05	3.53795	9.01937	2.90819	7.95504
			. 0.01	3.13114	10.79198	2.11528	8.74795
3 ALPHA	17.19921	0.500	14.80493	19.88341	14.62996	19.76847	
			. 0.100	11.68067	24.24959	10.93366	23.46477
			. 0.05	10.78849	25.79088	9.73335	24.66508
			. 0.01	9.17368	28.97543	7.38740	27.01103

Bibliography

- [1] Aroian, L. A, A study of Fisher's z distribution, *Annals of Mathematical Statistics*, 12, pp 429-448, 1939.
- [2] Ahrens, L.H., The lognormal distribution of the elements, *Geochimica et Cosmochimica Acta*, 6, pp 121-131, 1954.
- [3] Bhattacharyya, G. K. and Fries, A., On the Inverse Gaussian multiple regression and model checking procedures, Elsevier Science Publishers B.V. (North Holland), Editor Basu, A.P., pp 87-100, 1986.
- [4] Bhattacharyya, G. K. and Fries, A., Inverse Gaussian regression and accelerated life tests, *Mathematical Statistics*, Hayward, California, pp 101-118, 1982a.
- [5] Bhattacharyya, G. K. and Fries, A., Fatigue failure models Birnbaum Saunders VS. Inverse Gaussian, *IEEE Transactions on reliability*, Vol.R-31, No.5, pp 439-441, 1982.
- [6] Babu, G.J. and Chaubey, Y.P., Asymptotics and Bootstrap for Inverse Gaussian regression, *Ann.Inst.Statist.Math*, Vol.48, No.1, pp 75-88, 1996.
- [7] Beirlant, J., Goegebeur, Y., Verlaak, R. and Vynckier, P., Burr Regression and Portfolio segmentation, ELSEVIER Science B.V. *Insurance: Mathematics and Economics* 23, pp 231-250, 1998.
- [8] Bagdonavicius, V. and Nikulin, M., Semi parametric models in accelerated life testing, *Queen's papers in Pure and Applied Mathematics*, Kingston, Ontario, Canada, No.98, 1995.
- [9] Bagdonavicius, V. and Nikulin, M., Additive and Multiplicative semiparametric models in accelerated life testing and Survival analysis, *Queen's papers in Pure and Applied Mathematics*, Kingston, Ontario, Canada, Vol.108, 1998.
- [10] Banarjee, A. K. and Bhattacharyya G. K., A purchase Incidence Model With Inverse Gaussian Interpurchase Times, *Journal of Americal Statist. Association*, 71, pp 823-829, 1976.
- [11] Banarjee, A. *et al.*, Bayesian results for the IG distribution with application, *Technometrics*, 21, pp 823-829, 1983.

- [12] Bachelier, L., Theorie de la speculation, Ann. Sci. Ec. Norm. Super, 17, pp 21-86, 1900. the
- [13] Barton, D.E., The Probability distribution function of a sum of squares, Trabajos de Estadistica, 4, pp 199-207, 1953.
- [14] Bliss, C.I., The method of Probits, Science, 79, pp 38-39, 1934.
- [15] Cramer, H., Mathematical methods of Statistics, Princeton University press, Princeton, 1946.
- [16] Cramer, J.S., Econometric applications of maximum likelihood methods, Cambridge University press, Cambridge, London, 1986.
- [17] Cochran, W.G., Some difficulties in the statistical analysis of replicated experiments. Empire Journal of Experimental Agriculture, 6, pp 157-163, 1938.
- [18] Chayes, F., The lognormal distribution of the elements: a discussion, Geochimica et Cosmochimica Acta, 6, pp 119-120, 1954.
- [19] Chhikara, R.S. and Folks, J.L., The IG distribution as a life time model, Technometrics, 19, pp 461-468, 1974.
- [20] Chhikara, R.S. and Folks, J.L., Optimum test procedures for the mean of first passage time distribution in Brownian motion with positive drift, Technometrics, 18, pp 189-193, 1976.
- [21] Curtiss, J.H., On transformations used in the analysis of variance, Annals of Mathematical Statistics, 14, pp 107-122, 1943.
- [22] Chhikara, R. *et al.* Prediction limit for the IG distribution, Techno. 21 pp 319-324, 1982.
- [23] Draper, N.R. and Smith, H., Applied regression Analysis, John Wiley and Sons Inc. New York, 1966.
- [24] Dhrymes, P.J., On devising unbiased estimators for the parameters of the Cobb-Douglas production function, Econometrica, 30, pp 297-304, 1962.
- [25] Das, S.C., The fitting of truncated type III curves to daily rainfall data, Australian Journal of Physics, 7, pp 298-304, 1955.

- [26] Dabrowska, D.M., and Doksum, K.A., Partial likelihood in Transformations Models with Censored Data, *Scand. J. Statist.*, 15, pp 1-23, 1988.
- [27] Dugue, D. Application des propriétés de la limite au sens du calcul des probabilités à l'étude des diverses questions d'estimation, *J.Ecol. Poly.*, 3, No. 4, 305, 1937.
- [28] Eaton, W.W. and Whitmore, G.A., Length of stay as a stochastic process: a general approach and application to hospitalization for schizophrenia. *J. Math. Social*, 5, pp 273-292, 1977.
- [29] Ferrel, E.B., Control charts for log normal universes, *Industrial Quality Control*, 15, No.2, pp 4-6, 1958.
- [30] Gallant, A.R., *Nonlinear Statistical models*, John Wiley and Sons, New York, 1978.
- [31] Giordano, A.A. and Hsu, F.M., *Least square estimation with applications to digital signal processing*, John Wiley and Sons, New York, 1962.
- [32] Goon, A.M., Gupta, M.K. and Dasgupta, B., *An outline of Statistical theory (Volume-2)*. The World press private limited, Calcutta, India, 1987.
- [33] Gupta, S.C. and Kapoor, V.K., *Fundamentals of Mathematical Statistics*, Sultan Chand and Sons, Daryaganj, New Delhi-110002, 1970.
- [34] Greene, W.H., *Econometrics analysis*, Second edition, Maxwell Macmillan International Publishing Group, New York, 1993.
- [35] Galton, F., The geometric mean in vital and social statistics, *Proceeding of the Royal Society of London*, 29, pp 365-367, 1879.
- [36] Gaddum, J.H., Reports on biological standards, II: Methods of biological assay depending on a quantal response, special Report series, Medical research council, London, No. 183, 1933.
- [37] Goba, F.A., Bibliography on Thermal Aging of Electrical Insulation, *IEEE, Trans. on Electrical Insulation*, EI-4, pp 31-58, 1969.
- [38] Gibrat, R., *Les Inegalites Economiques*, Paris: Libraire du Recueil Sirey, 1931.
- [39] Gibrat, R., Une loi des répartitions, l'effet proportionné, *Bulletin de statistique, General*, France, 19, 469ff, 1930.

- [40] Grundy, P.M., The expected frequencies in a sample of an animal population in which the abundances of species are log normally distributed, Part-1, *Biometrika*, 3, 8, pp 427-434, 1951.
- [41] Grestein, G. *et al.*, Random walk models for the spike activity of a neuron. *Biophys. Jour.*, 4, pp 41-68, 1964.
- [42] Heyde, C.C., On a property of the lognormal distribution, *Journal of the Royal Statistical Society, Series B*, 25, pp 392-393, 1963
- [43] Hermanson, R.E. and Johnson, H.P., Generalized flood- frequency relationship, *Iowa state Journal of Science*, 41, pp 247-268, 1967.
- [44] Herdan, G.. The relation between the dictionary distribution and the occurrence distribution of word length and its importance for the study of quantitative linguistics, *Biometrika*, 45, pp 222-228, 1958.
- [45] Herdan, G., *The advanced theory of language as choice and chance*, New York, Springer, 1966.
- [46] Huzurbazar, V.S., The likelihood equation, consistency and maxima of the likelihood function, *Ann. Eugen. London*, 14, 185, 1948.
- [47] Johnson, N.L. and Kotz, S., *Continuous univariate distributions-I*, John Wiley and Sons, New York, 1970
- [48] Johnson, R.C.E. and Johnson, N.L., *Survival models and data analysis*, John Wiley and Sons, New York, 1980.
- [49] Kotz, S. and Neumann, J., On distribution of precipitation amounts for the periods of increasing length, *Journal of geophysical research*, 68, pp 3635-3641. Addendum: 69, 800-801, 1963.
- [50] Khamis, S.H., Incomplete Gamma function expansions of Statistical distribution functions, *Bulletin of the International Statistical Institute*, 37, pp 385-396, 1960.
- [51] Kalbfleisch, J.D. and Prentice, R.L., *The Statistical Analysis of failure time data*, John Wiley and Sons, New York, 1980.
- [52] Kapur, J.N. and Saxena, H.C., *Mathematical Statistics*, S.Chand and company Ltd. Ramnagar, New Delhi-110055, India, 1981.

- [53] Kolmogorov, A.N., Uber das logarithmisch normale verteilungs gesetzder Teilchen bei Zerstückelung, Doklady Akademii Nauk, SSSR, 31, pp 99-101, 1941.
- [54] Kartashov, G.D., Methods of Forced (Augmented) experiments (in Russian), Znaniye Press, Moscow, 1979.
- [55] Koopmans, L.H. Owen, O.B. and Rosenblatt, J.I., Confidence intervals for the coefficient of variation for the normal and lognormal distⁿ., Biometrika, 51, pp25-32, 1964.
- [56] Lehmann, E.L., Testing Statistical Hypothesis, John Wiley and Sons, New York, 1986.
- [57] Lawless, J.F., Statistical Models and Methods for Life time Data, John Wiley and Sons.. New York. 1982.
- [58] Lancaster, A., Stochastic model for the duration of a strike, J. Roy. Statist. Soc. ser. A., 135, pp 257-271, 1972
- [59] Mensah, R.D., Some aspects of survival analysis, M.Sc. thesis, Concordia University, Montreal, Quebec, Canada, 1981.
- [60] Miner, M.A., Cumulative Damages in Fatigue, Journal of Applied Mechanics, 12, pp A159-A164, 1945.
- [61] Miller, R.L. and Goldberg, E.D., The normal distribution in geochemistry, Geochimica et Cosmochimica Acta, 8, pp 53-62, 1955.
- [62] Morrison, J., The lognormal distribution in quality control, Applied Statistics, 7, pp 160-172, 1958.
- [63] Marcus, Allan H., Some exact distⁿ. in traffic noise theory, Adv. in appl. probab., 7, pp 593-606, 1975.
- [64] Neter, J., Wasserman, W., Kutner, M.H, Applied linear Statistical Models, IRWIN, Home Wood, IL60430, Boston, MA02116, 1974.
- [65] Nelson, W., Accelerated testing, Statistical models, Test plans and Data analysis, John Wiley and Sons, New York, 1990.
- [66] Nelson, W., Applied life data analysis, John Wiley and Sons, New York, 1981.

- [67] Nelson, W., Analysis of accelerated life test data, Part-1: The Arrhenius model and Graphical methods, IEEE Transactions on electrical insulation, Vol.EI-6, No.4, pp165-181, 1971.
- [68] Nelson, W., Analysis of accelerated life test data- Part-2: Numerical methods and test planning, IEEE Transactions on electrical insulation, Vol.EI-7, pp 36-55, 1972.
- [69] Nelson, W., Analysis of accelerated life test data- Part-3: Product comparisons and checks on the validity of the model and data, IEEE Transactions on electrical insulation, Vol.EI-7, pp 99-119, 1972.
- [70] Nelson, W., Graphical analysis of accelerated life test data with the Inverse Power Law model, IEEE transactions on reliability, Vol.R-21, No.1, pp 2-11, 1972.
- [71] Nelson, W., A survey of methods for planning and analyzing accelerated tests, IEEE transactions on electrical insulation Vol.EI-9, No.1, pp12-18, 1974.
- [72] Oldham, P.D., On estimating the arithmetic means of lognormally distributed populations, Biometrics, 21, pp 235-239, 1965.
- [73] Pearson, K. (Ed.), Tables of the Incomplete gamma function, London: H.N. stationary office, (since 1934, Cambridge University Press), 1922.
- [74] Pearson, K. (Ed.), On a criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, Philosophical Magazine, 5th series, 50, pp157-175, 1900.
- [75] Pearce, S.C., Lognormal distⁿ., Nature, London, 156, pp 747, 1945.
- [76] Prohorov, Yu. V., On the Lognormal distⁿ. in geo-chemistry, Teoriya Veroyatnostie i ee Primeneniya, 10, pp 184-187, 1963.
- [77] Rao, C.R., Linear Statistical Inference and its applications, Second edition, Wiley Eastern limited, New Delhi, India, 1984.
- [78] Rohn, W.B., Reliability prediction for complex systems, proceedings of the 5th National Symposium on Reliability and Quality Control in Electronics, pp 381-388, 1959.

- [79] Sedyakin, N.M., On one physical principle in reliability theory, *Techn., Cybernetics*, 3, pp 80-87, 1966.
- [80] Stuart, A. and Ord, J.K., *Kendall's Advanced theory of Statistics, Volume-2*, Oxford University press, New York, 1991.
- [81] Sprott, D.A. and Kalbfleisch, J.D., Examples of likelihoods and comparison with point estimates and large sample approximations, *American Statistical Association Journal*, pp 468-484, 1969.
- [82] Shenton, L.R. and Bowman, K.O., *Maximum likelihood estimation in small samples*, Charles Griffin and company Ltd. London and High Wycombe, 1977.
- [83] Singpurwalla, N.D., Inference from accelerated life tests using Arrhenius type re-parametrizations, *Technometrics*, Vol.15, No2, pp 289-299, 1972.
- [84] SAS/OR Technical Report: The NLP Procedure, SAS Institute Inc., SAS Campus Drive, Cary, NC 27513, USA, 1997.
- [85] Schrodinger, E., Zur Theorie der Fall-und steigversuche an Teilchenn mit Brownscher Bewegung, *physikalische zeitschrift*, 16, pp 289-295, 1915.
- [86] Tarikere, N.N., *Inverse Gaussian distribution and analysis of factorial experiments*, M.Sc. thesis, Concordia University, Montreal, Quebec, Canada, 1991.
- [87] Tweedie, M.C.K., Statistical properties of Inverse Gaussian distribution, *Annals of Mathematical Statistics*, 28, pp 362-377, 1957.
- [88] Tomlinson, R.C., A simple sequential procedure to test whether average conditions achieve a certain standard, *Applied Statistics*, 6, pp 198-207, 1957.
- [89] Tiku, M.L., Estimating the parameters of lognormal distribution from censored samples, *Journals of the American Statistical Association*, 63 pp 134-140, 1968.
- [90] Tiku, M.L. Approximating the general non-normal variance ratio sampling distribution, *Biometrika*, 51, pp 83-95, 1964.
- [91] Tiku, M.L., Laguerre series forms fo non-central chi-square and F-distributions, *Biometrika*, 52, pp 415-427, 1964.

- [92] Uven, M.J. Van, Logarithmic frequency distribution, proceeding of the Royal Academy of Science, Amsterdam, 19, pp 533-546, 1917.
- [93] Uven, M.J. Van, Skew frequency curves, Proceeding of the Royal Academy of Science, Amsterdam, 19, pp 670-684, 1917.
- [94] Wicksell, S.D., On the genetic theory of frequency Arkiv for Mathe-matik, Astronomi och Fysik, 12, pp 1-56, 1917.
- [95] Wise, M.E., Tracer dilution curves in cardiology and random walk and log-normal distribution. Acta Physiologica Pharmacologica Neerlandica, 14, pp 175-204, 1966.
- [96] Williams, C.B., The use of logarithms in the interpretation of certain entomological problems, Annals of Applied Biology, 24, pp 404-414, 1937.
- [97] Williams, C.B., A note on the statistical analysis of sentence length, Biometrika, 31, pp 356- 361, 1940.
- [98] Wu, Chuan-yi, The types of limit distributions for some terms of variational series, Scientia Sinica, 15, pp 749-762, 1966.
- [99] Wald, A., Sequential Analysis, John Wiley and Sons., New York, 1947.
- [100] Wasan, M.T. and Roy, L.K., Tables of Inverse Gaussian probabilities, Annals of Mathematical Statistics, 38, pp 299, 1967.
- [101] Welech, B.L., The significance of the difference between two means when the population variances are unequal, Biometrika, 29, pp 350-361, 1938.
- [102] Wasan, M.T., First passage time dist. of Brownian motion, Monograph, Dept. of Mathematics, Queen's University, Kingston, Ontario, 1968.
- [103] Weiss, G., Random walks and their applications, Amer. Scientist., 71, pp 65-71, 1983.
- [104] Whitmore, G.A., An IG model for labour turnover, J.R. Statist. Soc. series A, 142(A), pp 468-479, 1979.
- [105] Whitmore, G.A., A regression method for censored inverse gaussian data, Canada, J. Statist., 11, pp 305- 315, 1983.

- [106] Yurkowsky, W., Schafer, R.E. and Finkelstein, J.M., *Accelerated Testing Technology*, Rome Air Development center Tech. Rep. RADC-TR-67-420, Griffiss AFB, NY., 1967.