PHASE TYPE MODELLING
IN
RELIABILITY AND SURVIVAL ANALYSIS

Thesis submitted to the

University of Calicut

for the award of the degree of

DOCTOR OF PHILOSOPHY

under the Faculty of Science

by

ANEESH KUMAR, K.

DEPARTMENT OF STATISTICS

UNIVERSITY OF CALICUT

KERALA - 673 635

INDIA.

April 2007.

DECLARATION

I hereby declare that this thesis entitled **"Phase Type Modelling in Reliability and Survival Analysis"**, submitted to the University of Calicut, for the award of Doctor of Philosophy, is an independent work done by me under the guidance and supervision of Dr.M.Manoharan, Reader, Department of Statistics, University of Calicut.

I also declare that this thesis contains no material which has been accepted for the award of any other degree or diploma in any University and to the best of my knowledge and belief, it contains no material previously published by any other person, except where due reference is made in the text of the thesis.

Calicut University Campus,
11th April 2007.

Aneesh Kumar.K.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT

Dr. M. MANOHARAN, M.Sc., Ph.D.
Reader

CALICUT UNIVERSITY (P. O.)
KERALA, INDIA 673 635.
Phone : 0494 – 2401144 extn. 340 (O)
0494 – 2403403 (H)
Mob : 9447424043
Email : mano30@rediffmail.com

Date : 11th April 2007

CERTIFICATE

This is to certify that the work reported in this thesis entitled **PHASE TYPE MODELLING IN RELIABILITY AND SURVIVAL ANALYSIS** that is being submitted by Sri. Aneesh Kumar, K. for the award of Doctor of Philosophy, to the University of Calicut, is based on the bonafide research work carried out by him under my supervision and guidance in the Department of Statistics, University of Calicut. The results embodied in this thesis have not been included in any other thesis submitted previously for the award of any degree or diploma of any other university or institution.

Dr. M. Manoharan

ACKNOWLEDGMENT

First, I bow my head before my Lord who gave me enough care and power to complete this work. Next I would like to express my great indebtedness to my supervising teacher, Dr. M. Manoharan, Reader, Department of Statistics, University of Calicut, for his invaluable guidance, constructive criticism and encouragement throughout the work, without which the thesis could not have been materialized.

In this time I remember my teachers of the Department, Dr. N. Raju and Dr. K. Jayakumar for their constant motivation in my research work, I gratefully acknowledged their benevolence. Along with this, I wish to extend my sincere thanks to Dr. C. Chandran, Head, Department of Statistics, for his inspiring advice and tremendous support.

I am highly thankful to the Management, Principal and staff of Mahatma Gandhi College, Iritty. I bear in mind their goodness to provide me the facilities necessary for the research. Also I thank my parents, sisters, relatives, colleagues and friends for the numerous ways in which they helped me during the period of research.

I am very grateful to my wife Reshma for her constant support and encouragement.

I am very thankful to Mr.S. Suthagaran (Librarian) and the members of office staff, Department of Statistics for the limitless help they provided. I remember the Researchers, M.phil, M.Sc. students of the department who have made the tenure of research a pleasant experience with their encouragement and co-operation.

The financial support received from the UGC, Xth plan through FIP fellowship is gratefully acknowledged.

Last but not least, I express my hearty thanks to all those who helped me to complete the thesis in time.

Aneesh Kumar.K.

Contents

1 INTRODUCTION **5**

1.1 An Introduction 6

1.2 Phase Type Distribution - An overview 15

1.3 Reliability - An Introduction 18

1.4 Survival Analysis - An Introduction 22

1.5 An overview of the main contribution of the thesis 26

2 STRUCTURAL ASPECTS OF PHASE TYPE DISTRIBUTION **30**

2.1 Introduction 31

2.2 Discrete Phase Type (DPH) Distribution 32

2.3 Properties of DPH distribution 35

2.3.1 Analytical properties 35

2.3.2 Closure Properties 38

2.3.3 Associated Renewal Sequence 40

2.3.4 Renewal limit distributions 42

2.4 Continuous Phase Type (CPH) Distribution 43

2.5 Properties of CPH distribution 44

2.5.1 Analytical properties 45

2.5.2 Closure Results 45

2.5.3 Associated Renewal Process 48

2.5.4 Asymptotic Behaviour of CPH 51

2.6 Characterization Results 52

2.7 Characterization Result Under Poisson Shock Model 55

2.8 Fitting of Phase Type Distribution 58

3 ON WARRANTY WITH PREVENTIVE MAINTENANCE AND PHASE TYPE LIFE 62

3.1 Introduction 63

3.2 Determination of warranty period for a component 67

3.3 Warranty period for a series system 72

3.4 Numerical Illustration 78

3.5 Figures: 81

4 ON PREVENTIVE REPLACEMENT OF A DEVICE UNDER POISSON SHOCKS WITH DPH SHOCK PROBABILITIES 84

4.1 Introduction 85

4.1.1 Life of a device under Poisson shocks 86

4.1.2 Poisson shocks with Discrete Phase type shock probabilities 87

4.2 Optimal Replacement Age of the Device 89

4.2.1 Optimal replacement age by minimising long run cost per unit time for replacement 90

4.2.2 Optimal replacement age by considering a loss function . . 95

4.3 Figures: 102

5 MULTISTATE DETERIORATING SYSTEM SUBJECT TO POISSON SHOCKS 103

5.1 Introduction 104

5.2 The model 105

5.3 System under Poisson shocks with exponential shock magnitude . 110

5.4 Mathematical Illustration 110

5.5 System under Poisson shocks with unknown distribution for shock magnitudes 112

5.6 Figures: 115

6 MODELLING SEED GERMINATION - AN APPLICATION OF PHASE TYPE DISTRIBUTION 117

6.1 Introduction 118

6.2 The Model for Seed Germination 120

6.3 Estimation of Distributions 122

6.4 Experimental Validation 124

6.5 Tables and Figures: 130

A Concluding Remarks **134**

B Computer Codes **137**

B.1 Chapter 3 138

B.2 Chapter 4 139

B.3 Chapter 5 140

B.4 Chapter 6 141

C Bibliography **144**

INTRODUCTION

Aneesh Kumar K. “Phase type modelling in reliability and survival analysis”
Thesis. Department of Ststistics , University of Calicut, 2007

Chapter 1

INTRODUCTION

1.1 An Introduction

Phase type distributions are of great importance in modern probability and have gained widespread acceptance in recent years, because of its computational properties in applied stochastic modelling. A phase type distribution is the distribution of killing time in a finite state Markov chain. This versatile class is much applied in queuing theory, reliability and survival analysis.

When it is essential to have a more qualitative modelling to study real situations, like modelling deteriorating system, fluctuating failure rate, random environment etc., analytical approach to such models under general distributional assumption either fail or so complicated that they become intractable. By using phase type representation to those probability distributions, the essential features of the model can be reflected and useful information can be provided on its physical behaviour. Further more, there is no essential loss of generality in this set up. The class of phase type distributions is dense and hence any life distribution with positive support can be approximated arbitrarily close by a phase type distribution.

The common life distributions used in practice for reliability modelling are exponential and gamma. Practical experience over the past has suggested a much more flexible family of models. Phase type distributions are of proven utility in stochastic modelling but their application in reliability system is not all extensive, though several researchers, (Keilson(1975), Neuts and Meier(1981) and Assaf and

Levinskon(1982) etc.) have found various application of the theory of phase type distributions in reliability. Phase type distribution has the advantage that it (and its density, Laplace transform and all its moments) can be written in a closed form and thus, various probabilistic quantities of interest can be evaluated with relative ease. The closure and approximation results and the fact that phase type distribution is dense in the class of all distribution with support on $(0, \infty)$ make them useful in modelling different practical situations.

Several classes of life distributions used in reliability are closed under various reliability operations. Notably, the DFR class is closed under mixtures, the IFR class is closed under convolutions and IFRA class is closed both under convolutions and under the formation of binary coherent systems. Assaf and Levikson(1982) proved that the Phase type class is closed under all the above three operations. Dharmadhikari and Manoharan(1995) have shown that the class of phase type distributions is closed under the formation of multistate monotone system. Manoharan et.al.(1992) have obtained a preservation result of phase type distribution under Poisson shock models. These results are useful for evaluating reliability as well as various performance measures of a system. Chakravarthy(1983) considered a parallel system consisting of n identical components subject to failures. Failed components are repaired by a single repairman. Under the assumption that the failure times are exponential and the repair times are of phase type, it was shown that the busy period of the repairman, the density of the number of components repaired during a busy period and the stationary down time of the system are also of phase type. Chakravarthy (1987) studied the system of two machines in series with a buffer in between. The machines have exponential failure and repair times and the processing time is phase distributed. An algorithm for obtaining the steady-state probabilities and some system performance measures are presented. A complex priority redundant system with phase type distribution is analyzed by Gururajan and Bhat(1990), and the closed-form results for the reliability and availability of the system are provided. Chakravarthy et.al.(2001) used phase distribution for modelling the repair and service time of a k-out-of-n reliability system with an unreliable server. Time spent by a failed component in service, total time in the repair facility, vaccation time of the server, non-vaccation time of the server, and the time until failure of the system are all shown to be phase type. Barron et.al(2005) considered exponential life time and phase type repair time and proposed an algorithmic approach in analysis of R out of N system with several repairmen. They derived formulae for the point availability,the limiting availability, the distribution of the down time and the up time. Wei Li and Alfa(2005) considered a problem of optimal scheduling of a set of jobs on a single machine where the machine's uptime is assumed to have a general phase type with possible breakdowns and long repairs, and the processing time of the job are all stochastic. In a recent paper of Cui and Li(2007), an engineering system with structural as well as failure dependence among their components is discussed. Using the Markov process approach, a coherent system of components with multivariate phase type life distribution is introduced to model the components working in a common random environment. Apart from the above mentioned situations, there exists many other works in reliability exploiting nicely the properties of phase type distribution. In this thesis we bring together the properties of this versatile class of distribution and augment some more properties of it and use the distribution in various real life situations arising in reliability theory and survival analysis.

Post-sale service is an important element of new product sales. A warranty offers protection to buyers against early failures that might occur during the warranty period. Offering any warranty causes the manufacturer to incurs additional warranty servicing costs. For a product sold with warranty, the cost of servicing the warranty depends on the reliability of the product. There are many issues related to warranty. These includes selection of a warranty policy, warranty cost, reliability management policies, and many related areas and interaction between these. Murthy and Blischke(2001) describes various warranty policies and important aspects of warranty and reliability. Among the different warranty policies described, in FRW (Free Replacement Warranty) policy the manufacturer agrees to repair or provide replacements for failed items free of charge upto a time W from the time of the initial purchase; the warranty expires at time W after purchase. In PRW (Pro-Rata Warranty) policy the manufacturer agrees to refund a fraction of the purchase price should the item fail before time W from the time of the initial purchase. The buyer is not constrained to buy a replacement item. In RI (Reliability Improvement) policy the manufacturer agrees to repair or provide replacements free of charge for any failed parts or units until time W after purchase. In addition, the manufacturer guarantees the MTBF (Mean Time Before Failure) of the purchased equipment to be at least M. If the computed MTBF is less than M, the manufacturer will provide, at no cost to the buyer; (1) Engineering analysis to determine the cause of failure to meet the guaranteed MTBF requirement (2) Engineering change proposals (3) Modification of all existing units in accordance with approved engineering changes and (4) Consignment spares for buyer use until such time as it is shown that the MTBF is at least M. Nguyen and Murthy [1989], Jack and Dagpunar [1994] are some among who discussed warranty problems. The behaviour of lifetime is evidently an important aspect of warranty problems. It is of interest to learn more about how a warranty policy can be determined under a general lifetime distribution assumption. The use of phase type models seems to be quite worthwhile in such investigations as it happens to be in the present work.

There are several reasons for using the class of phase type distribution as a parametric model for failure time data. The class contains commonly used distributions like exponential, Erlang, sum of exponential and mixture of exponential. Then, when fitting a phase type distribution one get an automatic model selection within a large class of distributions.

In survival analysis when one considers the time for the occurrence of certain event, it may be noted that behind these events, there is an underlying process going through a set of stages which are only partially observed. For example, a disease goes through various stages of severity and which may be healed or may progress through gradual worsening to death. Modelling of seed germination is an important problem in plant science. Giving phase representation for the seed germination time, qualitative features of the various stages of germination process can be brought out and highly informative measures on the hazard rate of the stages be obtained.

Aalen(1995) describes the phase type methods for series model under acyclic Markov chain, where stage transition is allowed from one stage to the nearest stage only, with different transition intensities and plot its hazard function. It describes various kinds of phase type models, and connect them to problems in survival analysis. The model for incubation distribution of AIDS is an example for phase type model of acyclic type. In the model for the incubation time of AIDS, it considered two phases out of the five for developing to advanced HIV disease. In that stage treatment may be offered, described by a rate γ. The effect of treatment is to slow down the further progression of AIDS by factor θ. The incubation time in this model is the time it take to go from phase 1 to phase 5.

In biostatistical applications, the consideration of development of disease only in one direction will be unrealistic. For example, in the multistage model for cancer incidence one might allow cellular damage to be repaired. Hence it is more realistic to assume that the process moves back and forth between states even though absorption may eventually take place. Such models are so called "models with feed back". An appropriate modelling of such a situation is done in Aalen(1995) using phase type distributions. The paper also includes another example to illustrate fitting of phase type distribution to real data that concerns women having had a live-born first child, and the object is to analyse the time until their next birth, if any.

The length of stay of hospital patients can be effectively modelled using phase type distribution. An attempt in this direction is done in Faddy and McClean(2000). The phases are interpreted in terms of increased severity of any illness being treated. This leads to an identification of 'short stay', 'medium stay' and 'long stay' patients with the phase type distribution interpreted as a mixture of such components. Differential effects of two covariates, age of patient at admission and year of admission, are shown on the different phases of the distribution. Ahlstrom et.al.(1999) considered bivariate phase type distribution in a problem related with the parametric estimation procedure for relapse time distribution. Marshall and McClean(2003) modelled the length of stay of a group of elderly patients in a hospital using conditional phase type distribution. The model incorporates the use of Bayesian belief networks with Coxian phase type distributions; a special type of Markov model that describes the duration of stay in hospital as a process consisting of a sequence of latent phases.

The present thesis deals with phase type modelling, mainly in reliability and survival analysis. For phase type distribution many characterization results are derived in the literature. See for example, Aldous and Shepp(1987) and O'Cinneide(1991). In this thesis a characteristic property of phase type distribution is established based on the structure of a Poisson shock model.

A reliability and warranty problem of a product sold under warranty with general lifetime distribution is discussed in this thesis. An optimal warranty period for the product with preventive maintenances is suggested applying phase type distribution.

Preventive replacement of units is an important activity to avoid failures and associated cost or down time. The objective of the replacement problem is to find the optimal replacement age of the device that minimise the long run average cost or down time per replacement cycle. Many works like Cho and Parlar(1991), Suprasad Amari and Wes Fulton(2003), Berenguer et.al(2003) deals with the problems of preventive replacement. Here we find an optimal replacement age for a device working under Poisson shock using phase type distribution assumptions.

A vast majority of the reliability analysis assume that components and systems are in either of two states : functioning or failed. However in many real life situations we are actually able to distinguish among various 'levels of performance' for both system and components. For such cases, the existing dichotomous model is a gross over simplification and so models assuming degradable (multi-state) systems and components are certainly preferable. This topic was discussed earlier by Barlow and Wu(1978), El-Neweihi et.al (1978) etc. In this thesis we present the reliability analysis of a multistate deteriorating system subject to Poisson shock with general life distribution assumption.

Another area where phase type modelling is employed is that of seed germination. Seed germination is a complex biological process that is influenced by various environmental and genetic factors. The growth rate, hazard rate of seeds are of particular interest among plant scientists. Copeland(1988) describes various stages of seed germination. Shafili and Price(2001), Michael et.al(2004) considered modelling of seed germination time under various situation and distribution assumption. In this thesis we model seed germination time through phase type distribution assumption. Using phase type distribution, we modelled the life distribution of a seedling in its different growth stages and also the entire life, considering the data set including censored observations. The survival and hazard pattern of the seedling in its different growth stages are obtained.

In the following sections we narrate the earlier developments in the field of our current interest. In section 2, the development of phase type distribution and some of the important works on phase type distribution are discussed. An introduction to reliability theory is given in section 3. In section 4 an introduction to survival analysis is included. In the last section an overview of the main contribution of the thesis is provided.

1.2 Phase Type Distribution - An overview

Among the distributions of non-negative random variables, exponential distribution gained attention due to its simplicity and excellent analytical properties like memory less property. But this distribution is not closed under the basic convolution and mixing operation in probability. Noticing the limitations of exponential distribution, in the 20's of last century Erlang introduced the well known Erlang distribution through an ingenious method of stages (phases) which was later proved to be the main source of innovations in the field of stochastic modelling. Along this way Cox investigated a special series parallel network of stages and derived from it the famous Coxian Distribution. Through the use of complex transition probabilities Cox(1955) established the R-Distribution class characterised by rational Laplace-Stieltges Transforms(LSTs) and pointed out that in principle the distribution of any non-negative random variable may be approximated by a distribution in the R-Class. Jenson(1954) was the first to make the connection between the distribution of a non-negative random variable and a finite state Markov chain. As an extension of such ideas, Neuts(1975) initiated the phase type distribution theory. Phase type distribution retains many advantages of exponential distribution while overcoming some of its non-closure drawbacks. The monograph of Neuts(1981) established phase type distribution as a tool for unifying a variety of stochastic models and for constructing new models that yield to algorithmic analysis.

Neuts(1981) mentions the motivation for using this versatile class of probability distribution. In a large number of analytical investigation of stochastic models, it has been highly desirable to locate a Markov chain of sufficiently small dimension which would describe the given stochastic model completely computationally tractable and is free from the dimensionally problems often crops up while using the embedding techniques. Secondly, when one examines the models under non-exponential assumption, there arises the question related to the robustness or insensibility to the distributional assumption or to certain features of the model. Under phase type distribution assumption, the insensitivity results may be numerically examined and highly sensitive features of the model can be identified. The most established motivation for using it as a statistical model comes from their role as the computational vehicle of much of applied probability. Very often, problems which have an explicit solution assuming exponential distributions but not under general distributions are algorithmically tractable when one replaces the exponential distribution with a phase type distribution. Phase type distributions have been widely used in stochastic modelling in diverse fields. For typical examples, see Neuts(1981), Sengupta(1989), Asmussen(1992) in queuing theory; Assmussen and Rolski(1991) in insurance risk theory; Kao(1988), Lispky(1992) in renewal theory; Jeonson et.al.(1994) in computer science and so on. Latouche and Ramaswamy(1999) presents the basic ideas, method of phases and algorithms of matrix analytical theory.

The distribution of the time until absorption in a discrete parameter Markov chain is termed as Discrete Phase Type (DPH) distribution and time until absorption in a continuous parameter Markov chain is termed as Continuous Phase Type (CPH) distribution . A detailed discussion on phase type distribution and its properties is included in Chapter 2. Assaf et.al.(1984) discussed multivariate phase type distribution. Assaf et.al.(1985) proved that any phase type distribution can be represented as a proper mixture of two distinct phase type distributions. Shantikumar(1985) defined a bilateral phase type distribution. Soohan and Ramaswamy(2005) proved that bilateral phase type distribution is having many closure properties and is dense on the class of all distributions on real line. Following up Neut's idea, Shi Dinghua et.al.(1996) introduced the SPH distribution class, which describes on the absorption time in an infinite Markov chain. Shi Dinghua et.al.(2005) characterize the SPH class through the derivatives of the distribution functions. It is well known that DPH class has different properties with respect to the minimal coefficient of variation than CPH class, since the deterministic distribution (with zero coefficient of variation) is a member of DPH class. Motivated by the convenience in using the DPH family in applied stochastic modelling, Bobbio et.al.(2003) investigated more closely on the

DPH distributions and its acyclic subclass referred to as acyclic-DPH. Bobbio et. al.(2005) presents more on acyclic phase type (APH) distribution.

A statistical approach to estimation theory for phase type distribution is presented in Assmussen et.al.(1996). The idea of this approach is : the class of phase type distribution may for a fixed k transient states, be viewed as a multi parameter exponential family, provided the whole of the underlying absorbing Markov process is observed. Since the data in practice consist of i.i.d. replication of the absorption times $Y_1, Y_2, \ldots, Y_n$ of Y , there are incomplete observations and it is given to implement the EM algorithm. The program for implementation of the proposed algorithm is EMPHT-program by Haaggstrom et.al.(1992). The performance and the dynamics of the algorithm are illustrated in Assmussen et.al.(1996) by a sequence of fits of phase type distributions to three different theoretical distributions : Weibull, lognormal and Erlang distribution with feed back. As an extension to EMPHT-program, Olsson(1998) introduced a fitting tool, called EMpht programme to fit a CPH distribution. Bobbio and Telek(1994) also discussed the fitting of phase type distribution. A phase type fit tool recently introduced by Horvath and Telek(2002) can fit both CPH and DPH.

1.3 Reliability - An Introduction

Reliability considerations are important factors in any engineering system. The problems of failures, repairs and maintenances etc., are to be viewed seriously in the use of any equipments or systems., because the future performance or future life is entirely depending on such reliability characteristics which are not always deterministic due to the complexity, sophistication in nature. So to solve reliability problems it is highly essential to make use of stochastic models suitable for the situation.

Today design and production engineers discuss methods of improving reliability in design and production stages respectively. Testing engineers simulates the realistic environmental conditions. Statisticians produce the optimal experimental methods of life testing. Mathematicians construct mathematical methods of system reliability. Operations Research workers are concerned with the best methods of scheduling maintenance, replacement and repair of the equipment to obtain greater reliability matched with adequate availability and costs. All these investigations fall under the name of reliability studies, every year opens new vistas, but all the time probability concepts and statistical methods are used.

In reliability, we are mainly concerned with devices or system that fail at an unpredictable random age of $T > 0$. This random variable is assumed to have a distribution F, $F(t) = P(T \leq t)$, $t \in R$, with a density f. Then the reliability $R(t)$ of the system is defined as

$$P(T > t) = 1 - F(t) = \overline{F}(t) \tag{1.3.1}$$

This function is also know as the 'survival function'.

The failure rate λ is defined on the support of the distribution by

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)}$$

The failure rate $\lambda(t)$ measures the proneness to failure at time $'t'$ and we have the relation,

$$\overline{F}(t) = e^{-\int_0^t \lambda(s)ds} \quad \text{[see Aven and Jenson(1999)]}$$

Epestein and Sobel(1953) began the work in the field of life testing, which marked the beginning of wide spread research using exponential distribution. Until 1960 reliability was defined as the probability that an item will perform a required function under stated condition for a state of period of time. Research on coherent structure began with the paper by Birnbaum et.al.(1961). Barlow and Proschan(1965) stimulated the research work on the failure rate function and classes of life distributions defined in terms of this function. Barlow and Proschan(1975) can be viewed as first milestone in lifetime analysis, complex systems and maintenance models.

The notion of ageing is first discussed in Bryson and Siddique(1969). The concept of ageing for life distributions have been found very useful in reliability theory. Based on these notions, results may be derived concerning the behaviour of systems, bounds for survival functions, moment inequalities etc. It has been found useful for arriving at efficient algorithms for use in maintenance policies. Concerning the lifetime random variable T, the mean of T is denoted by μ.

For each t, with $\overline{F}(t) > 0$,

$$\overline{F}(x/t) = \frac{\overline{F}(t+x)}{\overline{F}(t)} = P[T \geq t + x/T \geq t]$$

represents the survival function of a unit of age t. The remaining (residual) life, at age t, is $E(T - t/T > t)$, which may be shown to be $\int_0^\infty \overline{F}(x/t)dx$. When $\lambda(t) = \frac{f(t)}{\overline{F}(t)}$ exits, $-log\overline{F}(x) = \int_0^x \lambda(t)dt$, represents the cumulative failure rate.

The following notions of positive (adverse effect) ageing are common in reliability theory. We say F is said to be

1. Increasing Failure Rate (IFR) if $\overline{F}(x/t)$ is decreasing in $0 \leq t \leq \infty$ for each $t \geq 0$. When the density exists, this is equivalent to $\lambda(t) = \frac{f(t)}{\overline{F}(t)}$ being non-decreasing in $t \geq 0$.

2. Increasing Failure Rate Average (IFRA) if $-\frac{log\overline{F}(x)}{x}$ is non-decreasing in x.

3. Decreasing Mean Residual Life(DMRL) if the mean remaining life function $\int_0^\infty \overline{F}(x/t)dx$ is non-increasing in x.

4. New Better than Used (NBU) if $\overline{F}(x/t) \leq \overline{F}(x)$ for $t \geq 0, x \geq 0$.

5. New Better than Used in Expectation (NBUE) if $\int_0^\infty \overline{F}(x/t)dx \leq \mu$ for $t \geq 0$.

6. Harmonically New Better than Used in Expectation (HNBUE) if

$$\int_0^\infty \overline{F}(x)dx \leq \mu e^{-\frac{\mu}{t}} \quad \text{for } t \geq 0.$$

7. L-Distribution if $\int_0^\infty e^{-st}\overline{F}(t)dt \geq \frac{\mu}{1+s}$ for $s \geq 0$.

8. New Better Than Used in Failure Rate (NBUFR) if $\lambda(t) \geq$ for $t \geq 0$.

9. New Better Than Used in Average Failure Rate (NBAFR) if

$$\lambda(0) \leq \frac{1}{t}\int_0^\infty \lambda(x)dx = -\frac{log\overline{F}(t)}{t}.$$

1.4 Survival Analysis - An Introduction

Survival analysis is a traditional statistical theme, but the recent surge of interest in this area is mainly due to its applications in biomedical sciences. In survival and event history analysis, one studies the time to occurrence of certain events. The field of survival analysis emerged in the 20^{th} century and experienced tremendous growth during the latter half of the century. The early efforts in development of survival analysis methodology were predominantly focussed as the estimation of the hazard function $\lambda(t)$ and the survival function $S(t)$. The developments in the field of survival analysis that have had the most profound impact on various application field are the Kaplan-Meier(1958) method for estimating the survival function, the log-rank statistic by Mantel(1966) for comparing two survival distributions and the Cox(1972) proportional hazard model for quantifying the effects of covariates on the survival time.

Survival data is a term used for data measuring the time to some event. In the simplest case, the event is death, but the term also covers other events, like occurrence of diseases, germination time of seeds etc. In industrial applications, it is typically time to failure of a unit or some component in a unit. In economics, it can be time acceptance of a job offer for an unemployed person. In demography, the event can be entering marriage.

Lifetime (survival time) data often come with a feature that creates problems called censoring in the analysis of the data. In broad sense it occurs when exact lifetime are known for only a portion of the individuals under study, the remaining lifetimes are known only to exceed certain values. Some times experiments are run over a fixed time period in such a way that an individual's lifetime will be known exactly only if it is less than some predetermined value. In such situation the data are said to be type I censored. The situation in which only the $'r'$ smallest observations in a random sample of n item are observed $(1 \leq r \leq n)$, the sample collected is said to be type II censored. It is to be noted that with type I censoring the number of exact lifetimes observed is random, incontrast to the case of type II censoring where it is fixed. Generalized type II censoring, random censoring etc., are some other types of censoring existing in literature. Lawless(1982) gives a detailed illustration on various types of censoring on survival (lifetime) data and also the parametric/non-parametric methodology for lifetime data analysis.

Some concepts in lifetime distribution:

Let T be a non-negative continuous random variable representing the lifetime of individuals in some population. Let $f(t)$ denote the probability density function (p.d.f.) of T and let the distribution function be

$$F(t) = Pr(T \leq t) = \int_{o}^{t} f(x)dx$$

and the survival function

$$S(t) = 1 - F(t)$$

In case of lifetimes of manufactured items, $S(t)$ has been referred to as the reliability function.

The p.d.f.(or p.m.f.) and the distribution and the survival function are common representation of a probability distribution, the hazard function (called the failure rate earlier) is a function which is particularly useful with lifetime distributions. They describes the way in which the instantaneous probability of death for an individual with time. Often, in applications there may be qualitative information about the hazard function, which can help in selecting a life distribution model. For example, there may be reason to restrict consideration to models with non decreasing hazard functions or with hazard function having some other well-defined characteristic. In short, the hazard function represents an aspect of a distribution that has direct physical meaning and that information about the nature of the hazard function is helpful in selecting a model.

The hazard function of a life distribution which specifies the instantaneous rate of death or failure at time t, given that the individual survives up till t, is denoted by $h(t)$ and is defined as,

$$h(t) = \lim_{\Delta t -> 0} Pr[\frac{t \leq T < t + \Delta t / T \geq t}{\Delta t}] \quad (1.4.1)$$

$$\Rightarrow h(t) = \frac{f(t)}{S(t)}, \text{ provided f(t) exists.} \quad (1.4.2)$$

It now follows that,

$$log_e S(t) = -\int_0^t h(x)dx.$$

$\int_0^t h(x)dx$ is the cumulative hazard function and is denoted by $H(t)$. Then,

$$S(t) = exp(-H(t))$$

$$\text{Also } f(t) = h(t)S(t)$$

$$= h(t)e^{-H(t)}.$$

When lifetime are grouped or when 'lifetime' refers to an integral number of cycles of some sort, it may be desired to treat T as a discrete random variable. Suppose T can taken on values $t_1, t_2, \ldots$ with $0 \leq t_1 \leq t_2 \leq \ldots$ and let the probability function be $P(t_j) = Pr(T = t_j)$ for $j = 1, 2, \ldots$. Then the survival function is,

$$S(t) = Pr(T \geq t) = \sum_{j=t_j \geq t} P(t_j)$$

$$\text{hazard function } h(t_j) = Pr(T = t_j / T \geq t_j)$$

$$= \frac{P(t_j)}{S(t_j)}, j = 1, 2, \ldots$$

As in continuous case, the probability, survivor, and the hazard functions give equivalent specifications of the distribution of T. Since $p_j = S(t_j) - S(t_{j+1})$,

$$h(t_j) = 1 - \frac{S(t_{j+1})}{S(t_j)}; j = 1, 2, \ldots$$

and,

$$S(t) = \prod_{j:t_j<t} [1 - h(t_j)].$$

Occasionally situations arises in which one would like T to have both discrete and continuous components. Special notation or definitions will not be introduced for such situations, which will be handled as they occur. No real difficulties are encountered with mixed distributions, especially if one works primarily with the survival function, which, as usual, is a monotone decreasing left-continuous function on $[0, \infty)$.

1.5 An overview of the main contribution of the thesis

The thesis is organised into six chapters and each chapter is divided into different sections. In the introductory chapter, a general introduction to the thesis is provided in the first section. Section 2 presents an overview on phase type distribution. A brief introduction of reliability and survival analysis is given in the next two sections. The last section provides an overview of the main contribution of the thesis.

Chapter 2 presents a detailed discussion on phase type distribution. Analytical properties, closure results etc., of the discrete and continuous form of phase type distribution is provided. Main characterisation results of the distribution in the literature is included in this chapter. Along with this we present a new characterization result on phase type distribution. We consider the life distribution H(t) of a device subject to shocks governed by a homogenous Poisson process. It is established that under certain mild conditions, $\{p_k\}$ - the probabilities that the device fails on the k^{th} shock, is a discrete phase-type (DPH) distribution is a necessary and sufficient condition for $H(t)$ to have a continuous phase-type (CPH) distribution. In particular, we prove that $H(t)$ follow exponential(μ) distribution, if and only if $\{p_k\}$ is a geometric distribution, under the homogenous Poisson(λ) shock model, when $\lambda > \mu$. This chapter ends with a description of fitting of phase type distribution using EM-algorithm.

In *chapter 3*, we consider a product with general life distribution that exhibits different life patterns at various places of installation according to the environmental conditions. A combination of Free Replacement Warranty(FRW) policy and Reliability Improvement (RI) policy is employed, while the product life time is modelled using Phase type distribution. Optimal warranty period with preventive maintenances is obtained and the method is illustrated through a numerical example.

Chapter 4 deals with an optimal preventive replacement policy for a device that sustains shocks those arrive according to a homogenous Poisson process. The discrete phase type distribution(DPH) is assumed for the failure (shock) probabilities. The DPH approximation for the failure probabilities could help to handle quite general situations. The problem assumes greater significance when loss due to preventive replacement is smaller than that due to the corrective replacement. For discrete phase type failure probabilities, the problem of finding the optimum value of the replacement time by minimising long run cost per unit time for replacement, and also the same problem by considering a particular loss function is discussed. A comparison of the above preventive replacement policies with the failure replacement is made proposing the gain due to age replacement.

In *chapter 5*, a multistate system is considered. The system is subject to homogenous Poisson shocks. The system gets deteriorated according to the magnitude of the shocks acted on it. The shock probabilities follow discrete phase type distribution and the parameters are determined using the probability distribution of shock magnitude. The life of the system is shown to have continuous phase type distribution. The reliability of the system is discussed.

A problem of modelling the seed germination time is discussed in *chapter 6.* The process of seed germination involves many stages,the end of which a mature plant suitable for plantation develops. The time required for the process and the sojourn time in each of its growth stages are modelled by phase type distribution. The growth rate and the hazard of the growing plant in its different growth stages are studied. The problem is illustrated using an experimental data. The EMpht program is used for the estimation procedure involved. The methodology suggested in this paper can be used for any seed germination process or similar ones.

All the mathematical illustrations appeared in this thesis are performed using MATLAB. The computer codes used for the purpose of each chapter are appended at the end of the thesis.

We conclude the thesis pointing out the salient features of our studies and scope for further work. A fairly comprehensive bibliography including all papers and books mentioned in this thesis is given at the end.

STRUCTURAL ASPECTS OF PHASE TYPE DISTRIBUTION

Chapter 2

STRUCTURAL ASPECTS OF PHASE TYPE DISTRIBUTION

2.1 Introduction

A non negative random variable (or its distribution) is said to be of phase type, if it is the time until absorption in a finite state, time homogeneous Markov chain. This family which we denote by PH, was introduced by Marcel F Neuts(1975) as a tool for unifying a variety of stochastic models and for constructing new models that yield algorithmic analysis. The idea of methods of phases is quite old, but the modern interest in this method, to a large extent, is due to Neuts, the main sources being his two books - (1981) ,(1989). The PH distributions are important in their own respects and have found useful in various fields such as queuing, reliability, survival analysis, branching process, telecommunication, computer science etc.

There are two parallel discussion of phase type distribution; one corresponds to absorption time in a discrete parameter Markov chain, which is termed as discrete phase type (DPH) distribution and and the other corresponds to absorption time in a continuous parameter Markov chain termed as continuous phase type (CPH) distribution. This chapter presents a detailed discussion on discrete and continuous phase type distributions and their properties. The characterization results on phase type distributions are reviewed and a new characterization of phase type distribution under Poisson shock model is given.

2.2 Discrete Phase Type (DPH) Distribution

Definition 2.2.1 *A probability density $\{p_k\}$ on the set of nonnegative integers is called a discrete phase type (DPH) distribution if it is the density of the time until absorption in a finite state Markov chain with stationary transition probability matrix given by*

$$P = \left[\begin{array}{c|c} T & \underline{T}^0 \\ \hline \underline{0} & 1 \end{array}\right] \qquad (2.2.1)$$

and initial probability vector $(\underline{\alpha}, \alpha_{m+1})$. Here T is an $m \times m$ substochastic matrix such that $T\underline{e} + \underline{T}^0 = e$ and $(I - T)$ is nonsingular.

The DPH density is given by

$$p_0 = \alpha_{m+1}, \quad p_n = \underline{\alpha} T^{n-1} \underline{T}^0, n \geq 1. \qquad (2.2.2)$$

The pair $(\underline{\alpha}, T)$ is called the representation of DPH and m is the order of the phase type distribution.

Examples of DPH distributions:

Example 2.2.1 *Let X follow a geometric distribution of the form*

$$P(X = x) = q^x p, \;\; x = 0, 1, 2, \ldots; \;\; 0 < p < 1; \;\; p + q = 1.$$

Here X is the waiting time for the first success. Consider a two state Markov *chain on the state space* $\{0,1\}$ *where* 1 *is an absorbing state. Then the distribution of time until absorption to state* 1 *starting from state* 0 *is the geometric distribution. Here*

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{c|c} q & p \\ \hline 0 & 1 \end{array}\right] \end{array}$$

where $\alpha = (1-p) = q$ *and* $T = q$

Thus geometric distribution is DPH $(\underline{\alpha}, \mathbf{T})$ *where* $\alpha = q$ *and* $T = q$

Example 2.2.2 *Generalized negative binomial distribution*

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ r-1 \\ r \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & \dots & r-1 & r \end{array} \\ \left[\begin{array}{ccccc|c} q_0 & p_0 & 0 & \dots & 0 & 0 \\ 0 & q_1 & p_1 & 0 & \dots & 0 \\ 0 & 0 & q_2 & p_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & q_{r-1} & p_{r-1} \\ \hline 0 & 0 & 0 & \dots & 0 & 1 \end{array}\right] \end{array}$$

where $p_i + q_i = 1$ *for* $i = 0, 1, 2, \dots, r-1.$

This is $DPH(\underline{\alpha}, \mathbf{T})_r$ *where*

$$\mathbf{T} = \begin{bmatrix} q_0 & p_0 & 0 & \cdots & 0 \\ 0 & q_1 & p_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & q_{r-1} \end{bmatrix} ; \underline{\alpha} = (1, 0, \ldots, 0)$$

Example 2.2.3

$$Let \;\; r_k = \left\{ \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right\}$$

Then

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \left[\begin{array}{ccc|c} 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \end{array}$$

This is $DPH(\underline{\alpha}, \mathbf{T})_2$ *where*

$$\mathbf{T} = \begin{bmatrix} 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} ; \;\; \underline{\alpha} = (1, 0, 0) \quad and \;\; \underline{\mathbf{T}}^0 = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{2} \\ 1 \end{bmatrix}$$

Example 2.2.4

$$Let\ r_k = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\}; \quad \mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \left[\begin{array}{ccc|c} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \end{array}$$

This is $DPH(\underline{\alpha}, \mathbf{T})_2$ *where*

$$\mathbf{T} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \quad \underline{\alpha} = (1,0,0) \ and \ \underline{\mathbf{T}}^0 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}$$

2.3 Properties of DPH distribution

The class of PH distributions possesses useful analytical properties as well as some interesting closure properties which makes it a good choice in the modelling of life times in various practical situations. Most of these results are well established in Neuts (1975, 1981). We discuss some important properties which have a direct bearing on the further development of the present work.

2.3.1 Analytical properties

Property 2.3.1 *The probability generating function(pgf) of DPH* $\{r_k, k \geq 0\}$, *with representation* $(\underline{\alpha}, \mathbf{T})$ *is given by* $R(z) = \alpha_{m+1} + z\underline{\alpha}[I - z\mathbf{T}]^{-1}\underline{\mathbf{T}}^0$, *for* $|\, z \,| \leq 1$

Note that this representation is valid, since for a DPH, $(I - \mathbf{T})^{-1}$ exists and $T^n \to 0$ elementwise. Therefore $(z\mathbf{T})^n \to 0$ elementwise for $\mid z \mid \leq 1$. Thus $(I - z\mathbf{T})^{-1}$ exists and is equal to $\Sigma_{k=0}^{\infty}(z\mathbf{T})^k$.

Property 2.3.2 *The j^{th} factorial moment of DPH $\{r_k, k \geq 0\}$ with representation $(\underline{\alpha}, \mathbf{T})$ is given by,*

$$\mu^{(j)} = j!\ \underline{\alpha}\mathbf{T}^{j-1}(I - \mathbf{T})^{-j}\mathbf{e} \quad for \quad j \geq 1$$

It follows that the mean of DPH

$$\mu_1' = \underline{\alpha}(I - \mathbf{T})^{-1}\mathbf{e}.$$

Also since $(I - \mathbf{T})^{-1}$ exists, it follows that all factorial moments of DPH exist and hence all raw moments exist.

Property 2.3.3 *Any probability density on a finite number of positive integers is of phase type.*

Property 2.3.4 *Every density obtained by shifting a density of phase type, an integer number of steps to the right is itself of phase type*

For an illustration, let us consider a DPH whose associated Markov chain is

$$\mathbf{P} = \begin{bmatrix} \mathbf{T} & \underline{\mathbf{T}}^0 \\ \mathbf{0} & 1 \end{bmatrix}$$

The probability density $\{p_k\}$ of DPH is given by,

$$p_0 = \alpha_{m+1}$$

$$p_k = \underline{\alpha}\mathbf{T}^{k-1}\underline{\mathbf{T}}^0, \; for \; k \geq 1$$

In order to shift the above density, 'r' places to the right, consider the representation, $(\underline{\beta}, \mathbf{S})$ of dimension$(m+r)$, given by,

$$\underline{\beta} = (1, 0, \ldots, 0) \text{ and } \mathbf{S} = \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ r-1 \\ r \\ r+1 \\ \vdots \\ r+m \end{matrix} \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & & & & \\ 0 & 0 & 1 & \ldots & 0 & & & & \\ 0 & 0 & 0 & \ldots & 0 & & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 & & & & \\ 0 & 0 & 0 & \ldots & 0 & \alpha_1 & \alpha_2 & \ldots & \alpha_m \\ & & & & & T_{11} & T_{12} & \ldots & T_{1m} \\ & & & & & \vdots & \vdots & \vdots & \vdots \\ & & & & & T_{m1} & T_{m2} & \ldots & T_{mm} \end{bmatrix}$$

It is clear that in an $(m+1)$ state Markov chain, if a state can be reached from another state, it can be reached in 'm' steps or less. Hence in a DPH, all of the terms $p_0, p_1, \ldots, p_m$ cannot be equal to zero. Suppose that $p_r, 1 \leq r \leq m$ is the first nonzero term of a PH density, then that density can be shifted upto 'r' steps to the left by changing the initial probability vector. The shifted densities are still of phase type. To illustrate, consider the density with

$$q_0 = \underline{\alpha}\mathbf{T}^{r-1}\underline{\mathbf{T}}^0 = p_r$$

$$\text{and} \quad q_k = p_{k+r} = \underline{\alpha}\mathbf{T}^r\mathbf{T}^{k-1}\underline{\mathbf{T}}^0, \quad \text{for} \quad k \geq 1$$

The density $\{q_k\}$ is of phase type with representation $(\underline{\beta}, \mathbf{T})$ with $\underline{\beta} = \underline{\alpha}\mathbf{T}^r$. It is easily verified that $p_j = 0, 0 \leq j \leq (r-1)$.This implies that $\underline{\alpha}\mathbf{T}^{r-1}\underline{\mathbf{T}}^0 + \underline{\alpha}\mathbf{T}^r\mathbf{e} = 1$. Thus the left shifts of the negative binomial density on $m, m+1, \ldots$ are therefore also of phase type.

2.3.2 Closure Properties

Property 2.3.5 *The convolution of a finite number of densities of phase type is itself of phase type.*

When we are considering two phase type distributions say $\{r_k\}$ and $\{s_k\}$, represented respectively by $(\underline{\alpha}, \mathbf{T})_m$ and $(\underline{\beta}, \mathbf{S})_n$, using the properties of pgf it is shown that, the convolution of $\{r_k\}$ and $\{s_k\}$ is phase type with parameters $(\underline{\gamma}, \mathbf{P})$, where

$$\underline{\gamma} = (\underline{\alpha}, \mathbf{0}, 0).$$

and,

$$\mathbf{P} = \begin{bmatrix} \mathbf{T} & \mathbf{T}^0\mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{S} & \underline{\mathbf{S}}^0 \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}$$

where $\mathbf{T}^0$ is a $m \times n$ matrix with n identical columns, each given by the vector $\underline{\mathbf{T}}^0$ and 'B' is $n \times n$ diagonal matrix with elements $(\beta_1, \beta_2, \ldots, \beta_n)$.

Property 2.3.6 *Any finite mixture of probability densities of phase type is itself of phase type.*

Let the mixing density be $(p_1, p_2, \ldots, p_\nu)$ and let the densities $r_k(h), h = 1, 2, \ldots, \nu$ be represented by the matrices $\mathbf{T}(h)$ and the initial probability vectors $\underline{\alpha}(h)$. The mixture $r_k = [\sum_{h=1}^{\nu} p_h r_k(h)]$ can be represented by the matrix,

$$\mathbf{P} = \left[\begin{array}{ccccc|c} \mathbf{T}(1) & 0 & 0 & \ldots & 0 & \underline{\mathbf{T}}^0(1) \\ 0 & \mathbf{T}(2) & 0 & \ldots & 0 & \underline{\mathbf{T}}^0(2) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \mathbf{T}(\nu) & \underline{\mathbf{T}}^0(\nu) \\ \hline 0 & 0 & 0 & \ldots & 0 & 1 \end{array}\right]$$

and the initial probability vector $\underline{\gamma} = (p_1\underline{\alpha}(1), p_2\underline{\alpha}(2), \ldots, p_\nu\underline{\alpha}(\nu), 0)$

But it is to be noted that an infinite mixture of phase type distributions are generally not of phase type.

Property 2.3.7 *Let $\{r_k\}$ be a density of phase type on positive integers defined by the vector $\underline{\alpha}$ and $m \times m$ matrix $\mathbf{T}$. Let $\{r_k^{(\nu)}\}$ be the ν fold convolution of the density $\{r_k\}$ with itself. Let $\{s_\nu\}$ be a density of phase type on the positive integers defined by the vector $\underline{\beta}$ and the $n \times n$ matrix $\mathbf{S}$. Define the probability density $\{v_k\}$ on the non-negative integers by,*

$$v_0 = \theta' \text{ and } v_k = \theta \sum_{\nu=1}^{\infty} S_\nu r_k^{(\nu)} \text{ for } k \geq 1 \tag{2.3.1}$$

where $\theta, \theta' > 0$ and $\theta + \theta' = 1$. Then $\{v_k\}$ is of phase type and is the density of the time till absorption in a Markov chain with $(m+1)$ states , the initial probability vector $(\theta \ \ \underline{\alpha} \otimes \underline{\beta}, \theta')$ and tpm

$$\mathbf{P}^* = \begin{bmatrix} \mathbf{T} \otimes I_n + (\mathbf{T}^0\mathbf{A}) \otimes \mathbf{S} & \underline{\mathbf{T}}^0 \otimes \underline{\mathbf{S}}^0 \\ \mathbf{0} & 1 \end{bmatrix} \tag{2.3.2}$$

where $\otimes$ is the Kronecker product operation, I_n is the $n \times n$ identity matrix, $\mathbf{T}^0$ is the $m \times m$ matrix with $(T^0)_{ij} = T_i^0$ and $\mathbf{A} = diag(\alpha_1, \alpha_2, \ldots, \alpha_m)$.

2.3.3 Associated Renewal Sequence

Consider a *DPH* $\{p_k, k \geq 0\}$ with representation $(\underline{\alpha}, \mathbf{T})$ with $\alpha_{m+1} = 0$. Let $\mathbf{Q} = \mathbf{T} + \mathbf{T}^0\mathbf{A}$. The matrix $\mathbf{Q}$ is clearly non negative. Also we have

$$\mathbf{Qe} = \mathbf{Te} + \mathbf{T}^0\mathbf{Ae} = \mathbf{Te} + \underline{\mathbf{T}}^0 \ \underline{\alpha}\mathbf{e} = \mathbf{Te} + \underline{\mathbf{T}}^0 = \mathbf{e}.$$

Therefore $\mathbf{Q}$ is a stochastic matrix. From a given Markov chain with tpm $\mathbf{P}$ of the form

$$\mathbf{P} = \begin{bmatrix} \mathbf{T} & \underline{\mathbf{T}}^0 \\ \mathbf{0} & 1 \end{bmatrix} \text{ and initial probability vector } \underline{\alpha}^* = (\underline{\alpha}, \alpha_{m+1})$$

with $\alpha_{m+1} = 0$, one can construct a Markov chain with tpm $\mathbf{Q}$ and initial probability vector $\underline{\alpha}$. When an absorption into the state $m+1$ occurs in the Markov chain with tpm $\mathbf{P}$, instantaneously restart the chain by performing an independent multinomial trial with the probabilities $\alpha_1, \alpha_2, \ldots, \alpha_m$. If we can denote the state of this new chain at time n by X_n we can see that $\{X_n, n \geq 0\}$ is a Markov chain.

Let us call the 'restarting' of the original Markov chain a 'renewal'. The time between the successive renewals form an ordinary renewal sequence in the usual sense. The distribution of the times between renewals is of phase type given by $\{p_k, k \geq 0\}$.

Remark 2.3.1 *In view of the above observation one can construct an extremely efficient algorithm to compute the 'renewal density' $\{h_n, n \geq 0\}$ of an ordinary renewal process in which the lifetimes distribution is of DPH.*

Remark 2.3.2 *Since h_n is the expected number of renewals at time 'n' and at any time 'n' atmost one renewal can occur, h_n is interpreted as the probability that at time n a renewal occurs.*

The following indicates the computation of the renewal density $\{h_n, n \geq 0\}$.

Theorem 2.3.1 *If the life time distribution $\{p_k, k \geq 0\}$ is DPH with representation $(\underline{\alpha}, \mathbf{T})$, then the associated renewal density $\{h_n, n \geq 0\}$ is given by,*

$$h_0 = 0$$

$$h_n = \underline{\alpha}(\mathbf{T} + \mathbf{T}^0\mathbf{A})^{n-1}\underline{\mathbf{T}}^0, \qquad n \geq 1.$$

Reducible T matrix:

We say that a square matrix T is reducible if by applying the same permutation to its rows and columns it can be brought into the form $\mathbf{T} = \left[\begin{array}{c|c} \mathbf{T}_{11} & 0 \\ \hline \mathbf{T}_{21} & \mathbf{T}_{22} \end{array}\right]$ where $\mathbf{T}_{11}$ and $\mathbf{T}_{22}$ are square matrices.

Property 2.3.8 *If the stochastic matrix $\mathbf{T} + \mathbf{T}^0\mathbf{A}$ is reducible, then the density $\{r_k\}$ may be represented by an initial probability vector $\underline{\beta}$ of dimension less than m and a matrix $\mathbf{B}$, which is a proper sub matrix of $\mathbf{T}$.*

2.3.4 Renewal limit distributions

Let $\{p_k, k \geq 0\}$ be a probability density on the nonnegative integers with $0 \leq p_0 < 1$, and with finite mean μ. Define the sequence, $\{q_k, k \geq 0\}$ by

$$\begin{aligned} q_k &= \frac{1}{\mu}\sum_{\nu=k}^{\infty} p_\nu \\ &= \frac{1}{\mu}[1 - \sum_{\nu<k} p_\nu], for \quad k \geq 1 \end{aligned}$$

The sequence $\{q_k, k \geq 1\}$ is a probability density, which will be referred to as the 'renewal limit density', associated with the density $\{p_k, k \geq 0\}$. The following theorem shows that the renewal limit density of a phase type distribution is also of phase type.

Note: It can be seen that the density $\{q_k\}$ is independent of p_0 for $0 \leq p_0 \leq 1$, and hence we may assume without loss of generality that $p_0 = 0$. Thus we can consider the density $\{r_k, k \geq 1\}$ has the representation $(\underline{\alpha}, \mathbf{T})$ with $\alpha_{m+1} = 0$. Also further assume that the stochastic matrix $\mathbf{T} + \mathbf{T}^0\mathbf{A}$, used in property (2.3.8) is irreducible. This implies that there exists a unique probability vector $\underline{\pi}$ with all positive components such that $\underline{\pi}\mathbf{T} + \mathbf{T}^0\mathbf{A} = \underline{\pi}, \quad \underline{\pi}\mathbf{e} = 1$

Theorem 2.3.2 *If the density $\{r_k\}$ is of phase type with representation $(\underline{\alpha}, \mathbf{T})$, then the corresponding renewal limit density $\{q_k\}$ is of phase type with representation $(\underline{\pi}, \mathbf{T})$.*

The theorem is established by showing that the pgf of the distribution $\{q_k, k \geq 1\}$ is the same as that of *DPH* with representation $(\underline{\pi}, \mathbf{T})$.

2.4 Continuous Phase Type (CPH) Distribution

Definition 2.4.1 *A distribution F on $[0, \infty)$ is a continuous phase type (CPH) if it is the distribution of time until absorption in a finite state Markov process* *with generator*

$$Q = \left[\begin{array}{c|c} T & \underline{T}^0 \\ \hline \underline{0} & 0 \end{array}\right] \quad \textit{and initial probability vector} \quad (\underline{\alpha}, \alpha_{m+1}).$$

$T = (T_{ij})$ is a nonsingular matrix of order m and satisfies $T_{ii} \leq 0, 1 \leq i \leq m, T_{ij} \geq 0$ for $i \neq j$. The distribution F is given by

$$F(x) = 1 - \underline{\alpha} \exp(Tx)\underline{e}, \quad x \geq 0 \tag{2.4.1}$$

and we say F has representation $(\alpha, T)_m$.

Examples of CPH distribution:

Example 2.4.1 *The exponential distribution $F(x) = 1 - e^{-\lambda x}, x \geq 0, \lambda > 0$ has CPH representation $(\underline{\alpha}, \mathbf{T})_m$ with $\underline{\alpha} = 1, \mathbf{T} = -\lambda$ and $m = 1$.*

Example 2.4.2 *The generalized Erlang distribution of order 'm' with parameters $\lambda_1, \lambda_2, \ldots, \lambda_m$ has the representation $\underline{\alpha} = (1, 0, \ldots, 0)$ and*

$$\mathbf{T} = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & \cdots & 0 \\ 0 & -\lambda_2 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \lambda_m \end{bmatrix}$$

2.5 Properties of CPH distribution

Here we discuss the important properties of CPH distribution. Many of them are highly useful in lifetime modelling.

2.5.1 Analytical properties

Property 2.5.1 *The Laplace-Stieltjes transform $f^*(s)$ of $F(x)$ is given by,*

$$f^*(s) = \alpha_{m+1} + \underline{\alpha}(sI - T)^{-1}\underline{\mathbf{T}}^0 \quad for \quad Re \quad s \geq 0$$

Property 2.5.2 *If $X \sim CPH(\underline{\alpha}, T)_m$, and $Y = aX$; then $Y \sim CPH(\underline{\alpha}, a^{-1}T)_m$*

Property 2.5.3 *The non central or raw moments of $F(x)$ are given by*

$$\mu_i' = \int_0^\infty y^i dF(y) = (-1)^i i! \underline{\alpha} T^{-i}\mathbf{e} \quad for \quad i \geq 1.$$

In particular the mean of CPH is given by :

$$\mu_1' = -\underline{\alpha} T^{-1}\mathbf{e}$$

and Variance of CPH is given by :

$$\begin{aligned}\mu_2 &= \mu_2' - (\mu_1')^2 \\ &= 2\underline{\alpha} T^{-2} e - (-\underline{\alpha} T^{-1} e)^2 \\ &= 2\underline{\alpha} T^{-2} e - (\underline{\alpha} T^{-1} e)^2\end{aligned}$$

2.5.2 Closure Results

Property 2.5.4 *If F(.) and G(.) are both continuous PH distributions with representations $(\underline{\alpha}, \mathbf{T})$ and $(\underline{\beta}, \mathbf{S})$ of orders m and n respectively, then their convolution $F * G(.)$ is a PH distribution with representation $(\underline{\gamma}, \mathbf{L})$ given by,*

$$\underline{\gamma} = [\underline{\alpha}, \alpha_{m+1}\underline{\beta}]$$

and

$$\mathbf{L} = \begin{bmatrix} \mathbf{T} & \underline{\mathbf{T}}^0\mathbf{B}^0 \\ \mathbf{0} & \mathbf{S} \end{bmatrix}$$

Property 2.5.5 *If $F(.)$ is a PH distribution with representation $(\underline{\alpha}, \mathbf{T})$, then,*

$$F^*(x) = \frac{1}{\mu_1'} \int_0^x [1 - F(u)]du, \ for\ x \geq 0$$

is a PH distribution with representation $(\underline{\pi}, \mathbf{T})$ where $\underline{\pi}$ is the stationary probability vector of the irreducible generator $\mathbf{Q}^* = \mathbf{T} + \mathbf{T}^0\mathbf{A}^0$

Property 2.5.6 *A finite mixture of CPH distributions is again a CPH distribution. If $(p_1, p_2, \ldots, p_k)$ is a mixing density and $F_i(.)$ has the representation $[\underline{\alpha}(i), \mathbf{T}(i)]$ of order $m_i + 1, 1 \leq i \leq k$, then the mixture $F(x) = \sum_{i=1}^k p_i F_i(x)$ has the representation,*

$$\underline{\alpha} = [p_1\underline{\alpha}(1), p_2\underline{\alpha}(2), \ldots, p_k\underline{\alpha}(k)] \ and\ \mathbf{T} = \begin{bmatrix} \mathbf{T}(1) & 0 & \cdots & 0 \\ 0 & \mathbf{T}(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{T}(k) \end{bmatrix}$$

of order $m = \sum_{i=1}^k m_i$

Property 2.5.7 *Let $\{S_\nu\}$ be a discrete PH density with representation $(\underline{\beta}, \mathbf{S})$ of order n and $F(.)$ a continuous PH distribution with representation $(\underline{\alpha}, \mathbf{T})$ of* *order m, then the mixture* $\sum_{\nu=0}^{\infty} S_\nu F^\nu(.)$ *of the successive convolutions of* $F(.)$ *is of phase type with representation* $(\underline{\gamma}, \mathbf{L})$ *of order* mn *, given by,*

$$\underline{\gamma} = \underline{\alpha} \otimes \underline{\beta}(I - \alpha_{m+1}\mathbf{S})^{-1} and$$

$$\mathbf{L} = \mathbf{T} \otimes I + (1 - \alpha_{m+1})\mathbf{T}^0\mathbf{A}^0 \otimes (I - \alpha_{m+1}\mathbf{S})^{-1}\mathbf{S} \qquad (2.5.1)$$

The height of the jump γ_{mn+1} *at zero and the vector* $\underline{\mathbf{L}}^0$ *are given by,*

$$\gamma_{mn+1} = \beta_{n+1} + \alpha_{m+1}\underline{\beta}(I - \alpha_{m+1}\mathbf{S})^{-1}\mathbf{S}^0 \text{ and } \underline{\mathbf{L}}^0 = \underline{\mathbf{T}}^0 \otimes (I - \alpha_{m+1}\mathbf{S})^{-1}\underline{\mathbf{S}}^0 \quad (2.5.2)$$

Property 2.5.8 *Let* X *and* Y *be independent random variables with PH distributions* $F(.)$ *and* $G(.)$ *having representations* $(\underline{\alpha}, \mathbf{T})_m$ *and* $(\underline{\beta}, \mathbf{S})_n$ *respectively. Let* $F_1(.)$ *and* $F_2(.)$ *be the distributions corresponding to* $max(X, Y)$ *and* $min(X, Y)$ *respectively, where* $F_1(.) = F(.)G(.)$ *and* $F_2(.) = [1 - F(.)][1 - G(.)]$*. Then* $F_1(.)$ *and* $F_2(.)$ *are also of phase type and are given as follows:*

$$F_1(.) \text{ has the representation } (\underline{\gamma}, \mathbf{L})_{mn+m+n} \text{ where}$$

$$\underline{\gamma} = [\underline{\alpha} \otimes \underline{\beta}, \beta_{n+1}\alpha, \alpha_{m+1}\beta] \text{ and}$$

$$\mathbf{L} = \begin{bmatrix} \mathbf{T} \otimes I + I \otimes \mathbf{S} & I \otimes \mathbf{S}^0 & \mathbf{T}^0 \otimes I \\ 0 & \mathbf{T} & 0 \\ 0 & 0 & \mathbf{S} \end{bmatrix}$$

and $F_2(.)$ *has the representation* $[\underline{\alpha} \otimes \underline{\beta}, \mathbf{T} \otimes I + I \otimes \mathbf{S}]$.

Now we have another interesting property. Let N be the number of arrivals in a Poisson process of rate λ during an interval $[0, X]$ where X is a random variable independent of the Poisson process. If $F(.)$ is the distribution of X, then,

$$a_k = \mathrm{P}[N = k] = \int_0^\infty e^{-\lambda u} \frac{(\lambda u)^k}{k!} dF(u), \;\; k \geq 0.$$

The pgf $A(z)$ of $\{a_k\}$ is given by $A(z) = f(\lambda - \lambda z)$ where $f(.)$ is the Laplace-Stieltjes transform of $F(.)$.The following property gives a method to compute the density $\{a_k\}$ without numerical integrations, if $F(.)$ is a PH distribution.

Property 2.5.9 *If $F(.)$ is a PH distribution with representation $(\underline{\alpha}, \mathbf{T})$ then $\{a_k\}$ is a discrete PH density with representation $(\underline{\beta}, \mathbf{S})$ given by*

$$\begin{aligned}
\underline{\beta} =& \lambda\alpha(\lambda I - \mathbf{T})^{-1} \\
\mathbf{S} =& \lambda(\lambda I - \mathbf{T})^{-1} and \\
\beta_{m+1} =& \alpha_{m+1} + \alpha(\lambda I - \mathbf{T})^{-1}\mathbf{T}^0 \\
\mathbf{S}^0 =& (\lambda I - \mathbf{T})^{-1}\mathbf{T}^0.
\end{aligned}$$

2.5.3 Associated Renewal Process

The probability distribution $F(t)$ may be used to construct a renewal process in the same way as in the discrete case.Define the quantities $Q_{ij} \; for \; 1 \leq i, j \leq m$ by,

$$Q_{ij} = T_{ij} + T_i^0 \alpha_j \;\; for \;\; i \neq j \tag{2.5.3}$$

$$Q_{ii} = -\sum_{j \neq i} T_{ij} - \sum_{j \neq i} T_i^0 \alpha_j = T_{ii} + T_i^0 \alpha_j.$$

so that $\mathbf{Q} = \mathbf{T} + \mathbf{T}^0 \mathbf{A}^0$ where $\mathbf{T}^0$ is an $m \times m$ matrix all of whose elements are $\underline{T}^0$ and $\mathbf{A}^0 = (1 - \alpha_{m+1})^{-1} diag(\alpha_1, \alpha_2, \ldots, \alpha_m)$.

$$\begin{aligned} Now \;\; \mathbf{T}^0 \mathbf{A}^0 &= \mathbf{T}^0 (1 - \alpha_{m+1})^{-1} diag(\alpha_1, \alpha_2, \ldots, \alpha_m) \\ &= (1 - \alpha_{m+1})^{-1} \mathbf{T}^0 \mathbf{A} \\ &= (1 - \alpha_{m+1})^{-1} \underline{\mathbf{T}}^0 \underline{\alpha} \end{aligned}$$

The matrix $\mathbf{Q}$ is of considerable importance. It is the infinitesimal generator of the phase type renewal process which can be obtained by restarting the Markov chain defined in Equation (2.4.1) instantaneously after each absorption (renewal) by performing a multinomial trial with probabilities $\alpha_1, \alpha_2, \ldots, \alpha_{m+1}$ and outcomes $1, 2, \ldots, m+1$. The times between successive renewals of such a renewal process is the CPH $F(.)$.

Consider the m state Markov chain with infinitesimal generator $\mathbf{Q}$. Suppose that upon absorption into state $m+1$, we instantaneously perform independent multinomial trials with probabilities $\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}$ until one of the alternatives $1, 2, \ldots, m$ occurs. Restarting the process $\mathbf{Q}$ in the corresponding state, let us consider the time of the next absorption and repeat the same procedure. It is easy to see that by continuing this procedure indefinitely, we will get a new Markov process in which state $m+1$ is an instantaneous state. When that process enters the instantaneous state, the probability that it stays there for $r \geq 1$ transitions is given by, $(1-\alpha_{m+1})\alpha_{m+1}^{r-1}$ and the number of visits to the instantaneous state are independent with the common geometric distribution.

If we further consider the Markov process in which the path functions are right-hand continuous,we will obtain a Markov process on $1, 2, \ldots, m$ with infinitesimal generator, $\mathbf{Q}^* = \mathbf{T} + \mathbf{T}^0\mathbf{A}^0$, where $\mathbf{T}^0$ is a $m \times m$ matrix with identical columns T^0 and $\mathbf{A}^0 = (1-\alpha_{m+1})^{-1}diag(\alpha_1, \alpha_2, \ldots, \alpha_m)$. Thus $\mathbf{T}^0\mathbf{A}^0 = (1-\alpha_{m+1})^{-1}T^0\alpha$, provided $\alpha_{m+1} \neq 1$.

It can be seen that the successive visits to the instantaneous state form a renewal process with underlying distribution $F(.)$ given by

$$F(x) = 1 - \alpha exp(\mathbf{T}x)\mathbf{e}, x \geq 0.$$

The point process so obtained is a PH renewal process .

Property 2.5.10 *Let $S(z)$ be the pgf of a density of phase type on the non-negative integers, represented by,*

$$S(z) = \beta_{n+1} + \beta z(I - zS)^{-1}\mathbf{S}^0, \tag{2.5.4}$$

and let $R^*(s)$ denote the Laplace-Stieltjes transform of a phase type density on $[0, \infty)$, represented by,

$$R^*(s) = \alpha_{m+1} + \alpha(sI - T)^{-1}\mathbf{T}^0, \tag{2.5.5}$$

then the mixture with Laplace-Stiletjes transform

$$V^*(s) = S[R^*(s)], \tag{2.5.6}$$

is of phase type with representation $(\underline{\gamma}, \mathbf{L})$, *where*

$$\underline{\gamma} = \underline{\alpha} \otimes \underline{\beta}(I_n - \alpha_{m+1}\mathbf{S})^{-1} \tag{2.5.7}$$

$$\mathbf{L} = \mathbf{T} \otimes I_n + \mathbf{T}^0 A \otimes (I_n - \alpha_{m+1}\mathbf{S})^{-1}\mathbf{S}. \tag{2.5.8}$$

2.5.4 Asymptotic Behaviour of CPH

In this section, a number of simple results on the asymptotically exponential nature of PH distributions is cited. These results are useful in the computation of the waiting time distribution for a number of queues.

Definition 2.5.1 *A probability distribution* $F(.)$ *is asymptotically exponential if and only if for some* $K > 0 and\ \eta > 0$,

$$1 - F(x) = Kexp(-\eta x) + o(exp(-\eta x)), as\ x \longrightarrow \infty. \tag{2.5.9}$$

Theorem 2.5.1 *If the matrix* $\mathbf{T}$ *is irreducible, any* PH *distribution* $F(.)$ *with representation* $(\underline{\alpha}, \mathbf{T})$ *is asymptotically exponential and* $-\eta$, *is the eigen value with largest real part of* T. *The constant* $K = \alpha\nu$, *where* ν *is the positive right eigen vector of* $\mathbf{T}$, *corresponding to* $-\eta$,*uniquely determined by the requirements that* $\mathbf{uv} = \mathbf{ue} = 1$, *where* $\mathbf{u}$ *is a left eigen vector of* $\mathbf{T}$, *associated with* $-\eta$.

Corollary 2.5.1 *A geometric mixture of the convolution powers of a* PH *distribution is asymptotically exponential.*

2.6 Characterization Results

The characterization of phase type distribution has been attempted by many researchers in the past, specifically through its structural property and variability through the concepts of coefficient of variation or that of majorization. Notable among them are the following.

Theorem 2.6.1 *A distribution on* $[0, \infty)$ *is phase type if and only if it is either the point mass at zero or*

1. *it has a continuous positive density on* $(0, \infty)$ *and*

2. *it has rational Laplace-Stieltjes transform with a unique pole of maximal real part.*

A discrete-time version of this is due to Soittola (1976) and Katyama et.al(1978). The class of phase type distributions represents the natural family to which Erlang's method of stages extends. So it is of intereset to have an idea of variability for a phase type distribution of a given order. The Erlang distribution $E(n, \lambda)$ is PH distribution of order n. With regard to comparison in terms of coefficient of variations, Aldous and Shepp (1987) have proved the following theorem.

Theorem 2.6.2 *The coefficient of variations of an order* n PH*-distribution is atleast* $\frac{1}{\sqrt{n}}$ *and the order* n *Erlang distribution is the only one to attain the bond*

The discrete counterpart of the above result is provided by Telek (2000). Another way of comparing variability is through majorization, the concept of which we present below for better clarity of ideas.

Another way of comparing variability is through the following concept of majorization introduced by Hardy et.al(1952). A more recent treatment of majorization may be found in Marshall and Olkin(1979)

Definition 2.6.1 *Let μ and ν be two probability measures on $\boldsymbol{R}^n$ with finite means. We say that ν majorizes μ if*

$$\int f d\nu \geq \int f d\mu \quad \textit{for all convex functions } f.$$

It is written as $\mu < \nu$. It implies, in particular the means of μ and ν are equal. This stochastic ordering has been found useful for comparing variability among life distributions.

Theorem 2.6.3 *A PH-distribution with an order n representation majorizes the order n Erlang distribution of the same mean.*

Remark 2.6.1 *By means of a tricky example O'Cinneide(1991) also establishes that the above Erlang distribution is not least variable among general distributions with rational transforms and hence it suggests that Aldous and Shepp's result hings on the Markov property and on the order of PH-distribution, rather than the more elementary quantity-its degree.*

Maier and O'Cinneidie(1992) characterises the class of continuous and discrete phase type distributions as they are closed under convolutions, mixtures and under the 'unary geometric' operation. They proved that the continuous classes are the smallest family of distributions that is closed under these operations and contains all exponential distributions and the point mass at zero. An analogous result holds for the discrete distribution also.

O'Cinneidie (1999) also put forward conjectures concerning phase type distributions. Renhai (2002) proved the steepest increase conjucture, which states that "for any phase type density $f(t)$, of order n, $\frac{f(t)}{t^{n-1}}$ is increasing for $t > 0$".

Shi Dinghua et al.(2005) presented an extension for the phase type distribution named as SPH distribution. It is the distribution of the absorption time in an infinite state Markov chain. Phase type distribution is given as a subclass of SPH. Shi Dinghua et al.(2005) characterized the SPH class through the derivations of the distribution function and further proved that the infinite mixture of SPH distributions is also an SPH, but it is true only in case of finite mixture for the phase type distributions.

2.7 Characterization Result Under Poisson Shock Model

In this section a new characterization of phase type distribution under Poisson shock model is given. The result of exponential distribution then follows as a special case. Here we consider the life distribution H(t) of a device subject to shocks governed by a homogenous Poisson process.

Consider a device subject to shocks occurring randomly over time according to a counting process $N = \{N(t), t \geq 0\}$. Let p_k be the probability that the device fails on the k^{th} shock, and $\overline{P_k} = \sum_{n=k+1}^{\infty} p_n$ be the probability that the device survives k shocks, $k = 0, 1, 2 \ldots$ where $1 = \overline{P_0} \geq \overline{P_1} \geq \overline{P_2} \geq \ldots$. Then it follows that the survival function $\overline{H}(t)$ of the device is given by

$$\overline{H}(t) = \sum_{k=0}^{\infty} P\{N(t) = k\}\overline{P_k} \tag{2.7.1}$$

When the shock probability $\{p_k\}$ has a specified property and that carried over to the life distribution $H(t)$, we say that the property is preserved under the above Poisson shock model.

The preservation results of various ageing properties like increasing failure rate (IFR), increasing failure rage average (IFRA), decreasing mean residual life (DMRL), new better than used(NBU), new better than used in expectation(NBUE), harmonic new better than used in expectation(HNBUE), when N is a homogenous/non-homogenous Poisson process or more general counting process are now well known.

Manoharan et.al(1992) presented the preservation theorem of phase type distribution under finite mixture of Poisson processes and established the relationship between the mean value of $\{p_k\}$ and $H(t)$. Now we supplement this result by proving that phase type distribution of the shock probability is a necessary condition for the life distribution to have phase type distribution under certain mild conditions on the underlying parameters.

Consider a device subject to Poisson shocks and getting deteriorated. The shock probabilities has a discrete phase type representation, then the life distribution of the system $H(t)$ has a continuous phase type distribution. (See Manoharan et. al (1992)). Now consider a device with the lifetime following continuous phase type distribution $CPH(\alpha, T)$ subject to Poisson(λ) shocks. Then, it is proved that the shock probabilities of the system follow discrete phase type (DPH) distribution under some restrictions on λ

Since the life of the device $H(t)$ follows $CPH(\alpha, T)_m$, where $T = [T_{ij}]$ then for any $\lambda \geq |T_{ij}|$ for all i,j ; there exists a stochastic matrix T^* such that,

$$T = \lambda(T^* - I) \tag{2.7.2}$$

$$ie, T^* = \frac{T}{\lambda} + I \tag{2.7.3}$$

where T^* satisfies the restrictions imposed on the transition probability matrix for a discrete phase type distribution and I is an identity matrix. Then we can represent $H(t) \sim CPH(\alpha, \lambda(T^* - I))$. Hence the survival function of the device is

$$\overline{H}(t) = \alpha e^{\lambda(T^*-I)}\underline{e} \tag{2.7.4}$$

where, $\underline{e} = (1, 1, \ldots 1)$. Again since the shock arrival is Poisson(λ), the survival function of the device can be expressed as

$$\overline{H}(t) = \sum_{k=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} \alpha(T^*)^k \underline{e} \tag{2.7.5}$$

which is the expression of the survival function of a device under homogenous Poisson shock with $DPH(\alpha, T^*)$ shock probabilities. Thus we proved the following result.

Theorem 2.7.1 *The life distribution $H(t)$ of the device under homogenous Poisson(λ) shock model is $CPH(\alpha, T)$ if, and only if, the shock probabilities $\{p_k\}$ is $DPH(\alpha, T^*)$, when $\lambda \geq |T_{ij}|$ for all i,j . Further $T = \lambda(T^* - I)$.*

- **A special case:**

Consider a device with exponential(μ) life, which can be expressed in CPH form as $CPH(\alpha, T)$,where $\alpha = (1)$ and $T = [-\mu]$

Then for any $\lambda > \mu$, we can find an element q, where $0 < q < 1$, such that ,

$$-\mu = \lambda(q - 1)$$

Hence here the expression of survival function of the device with life $CPH(\alpha, T)$,

$$\overline{H}(t) = \alpha e^{Tt}\underline{e} \tag{2.7.6}$$

can be expressed as

$$\overline{H}(t) = 1 \times e^{-\mu t} \times 1 \tag{2.7.7}$$

that is

$$\overline{H}(t) = e^{-\lambda(1-q)t} \tag{2.7.8}$$

which can be written as,

$$\overline{H}(t) = \sum_{k=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} q^k \tag{2.7.9}$$

this in the form of the survival function of a system under Poisson(λ) shock and DPH shock probabilities with parameter $\alpha = 1$ and $T^* = q$

This implies that for a device under Poisson shock, with life $CPH(\alpha, T)$,where $\alpha = (1)$ and $T = [-\mu]$, the shock probabilities $\{p_k\}$ follows $DPH(\alpha, T^*)$, where $\alpha = (1)$ and $T^* = q$, when $\lambda > \mu$, which is the geometric distribution with failure probability q.

2.8 Fitting of Phase Type Distribution

In order to fit a phase type distribution to a sample $\mathbf{y} = (\mathbf{y_1}, \mathbf{y_2}, \ldots \mathbf{y_n})$ an EM-algorithm can be used. The EM- algorithm is an iterative method to calculate maximum likelihood estimates. It is often used in situations where the observed data are incomplete in some sense, or, to put it another way, can be viewed as partial observations of a larger experiment than the experiment actually performed. For a general account of the EM-algorithm and its properties, one may refer to Dempster et.al.(1997) and Wu(1983).

The EM-algorithm for fitting phase type distribution was presented in detail in Asmussen(1996). Suppose that we want to fit a phase type distribution with representation $(\alpha, T)_m$ to a sample $y = (y_1, y_2, \ldots, y_n)$. The idea behind the EM-algorithm is to regard each observation y_i as an incomplete observation of the underlying Markov process $X_i(t)$. A sample of n complete observations of Markov processes on the intervals $(0, y_1], \ldots, (0, y_n]$ respectevely, has the following likelihood function

$$L(\alpha, T) = \prod_{i=1}^{m} \alpha_i^{B_i} \prod_{i=1}^{m} exp\{t_{ii} Z_i\} \prod_{i=1}^{m} \prod_{j=0;j\neq i}^{m} t_{ij}^{N_{ij}} \tag{2.8.1}$$

where,

B_i =the number of Markov processes starting in state i, i = 1,2,...,m.

Z_i =the total time spent in state i, i = 1,2,...,m

N_{ij} =the total number of jumps from state i to state j, for $i \neq j$,

i = 1,2,...,m and j = 1,2,...,m,m+1

The maximum likelihood estimates of the elements of α and T , based on a sample of n complete observations of the Markov process, are calculated as

$$\hat{\alpha_i} = \frac{B_i}{n}; \; \hat{t_{ij}} = \frac{N_{ij}}{Z_i}; \; \hat{t_i} = \frac{N_{i0}}{Z_i}; \; \hat{t_{ii}} = -(\hat{t_i} + \sum_{j=1;j\neq i}^{m} \hat{t_{ij}}), i,j = 1,2,\ldots,m. \quad (2.8.2)$$

It may be noted that the distribution of such a sample of Markov process is a member of a curved multi parameter exponential family with sufficient statistic

$$S = ((B_i)_{i=1,2,\ldots,m}, (Z_i)_{i=1,2,\ldots,m}, (N_{ij})_{i=1,2,\ldots,m,j=1,2,\ldots,m;\, i\neq j}).$$

However, the phase type distributed sample **y** does not include observed values of the sufficient statistic S, and this is where the EM-algorithm comes in.

In the so-called E-step, the first step in each iteration of the algorithm, the conditional expectation of the sufficient statistics, given the observed sample **y** and the current estimates of the parameters, is calculated. In the second step, the M-step, new estimates of the parameters are calculated by using the conditional expectation of the sufficient statistic as if it was the observed value. Thus, let $(\alpha, T)^{(k-1)}$ denote the estimate number (k-1) of (α, T), then iteration number k of the algorithm is:

E-step: Calculate

$$B_i^{(k)} = \sum_{\nu=1}^{n} E[B_i^{(\nu)}/y_\nu; (\alpha, T)^{(k-1)}], \quad \text{i=1,2,...,m.}$$

$$Z_i^{(k)} = \sum_{\nu=1}^{n} E[Z_i^{(\nu)}/y_\nu; (\alpha, T)^{(k-1)}], \quad \text{i=1,2,...,m.}$$

$$N_{ij}^{(k)} = \sum_{\nu=1}^{n} E[N_{ij}^{(\nu)}/y_\nu; (\alpha, T)^{(k-1)}], \quad \text{i=1,2,...,m, j=1,2,...,m,m+1, } i \neq j$$

Next we have,

M-step: The new estimators are given by,

$$\alpha_i^{(k)} = \frac{B_i^{(k)}}{n};\ t_{ij}^{(k)} = \frac{N_{ij}^{(k)}}{Z_i^{(k)}};\ t_i^{(k)} = \frac{N_{i0}^{(k)}}{Z_i^{(k)}};\ t_{ii}^{(k)} = -(t_i^{(k)} + \sum_{j=1;j\neq i}^{m} t_{ij}^{(k)}).$$

(Observe that in the E-step the single statistic in S are written as sums over the sample, eg. $B_i = \sum_{\nu=1}^{n} B_i^{(\nu)}$, where $B_i^{(\nu)}$ is the contribution to B_i from $X_i(t)$.)

The likelihood functions of $\mathbf{y}$ increases in every iteration of the algorithm until it reaches a stationary value. It is recommended to run the algorithm several times, starting with different initial values of the parameters.

While fitting phase type distribution to survival data, it may require to handle censored data also. Olsson(1996) presents a way to estimate parameters of phase type distribution of fixed order from cersored samples, using EM algorithm. Two types of censored data-right censored and interval censored-are considered. The performance of the algorithm is illustrated with example of estimated phase type distributions from both simulated and real data sets. As an implementation of the EM algorithm presented by Asmussen et.al.(1996) and Olsson(1996), Olsson(1998) introduced a C language based program called EMpht-programme for fitting phase type distribution. It can be used either to fit a phase type distribution to a sample or make a phase type approximation of another continuous distribution. EMpht-programme is an extension of the EMPHT-programme of Haaggstrom et.al.(1992). The main difference being that EMpht can handle samples which contains right/interval censored observations.

ON WARRANTY WITH PREVENTIVE MAINTENANCE AND PHASE TYPE LIFE

Aneesh Kumar K. “Phase type modelling in reliability and survival analysis” Thesis. Department of Ststistics , University of Calicut, 2007

Chapter 3

ON WARRANTY WITH PREVENTIVE MAINTENANCE AND PHASE TYPE LIFE

3.1 Introduction

Warranty has been a part of trade since ancient times, however the modern notion of warranty is a post-industrial concept. It is a contract between the consumer and the manufacturer, which is entered into upon sale of the product. Warranties are an integral part of nearly all consumers and commercial and many government transactions that involve product purchases. In such transactions, warranties serve somewhat different purpose for buyer and seller.

To face the competition in the field of marketing, it is better for a manufacturer to introduce an attractive warranty policy. By preventive maintenances, one can offer long warranty periods for the component with a desired level of reliability. Designing a proper warranty program has become an important marketing tool now, especially to advertise the quality of the product. Of course, offering warranty usually results in additional costs to the manufacturers. The warranty servicing cost depends on the warranty length, item reliability and other costs. By properly designing the warranty, the manufacturer can increase the sales and also the market share.

Often warranty specifies the commitment of the manufacturer to replace or repair a failed item within the predetermined warranty period. The free replacement warranty policy under which the manufacturer agrees to repair or provide replacements for failed item free of charge during the warranty period; the pro-rata warranty policy according to which the replacements are provided under pro-rated cost to the consumer and their combinations are the most commonly used warranty policies. There are several aspects of warranty and many disciplines involved in the analysis of these issues. One may refer to Murthy and Blischke (2001) for an excellent composition on various topics that link warranty and reliability.

Maintenance policies during warranty are analysed by several authors (see Nguyen and Murthy (1989), Jack and Dagpunar(1994) among all) in which a one dimensional warranty characterised by an interval called the warranty period is considered. The case of two-dimensional warranty is discussed in several papers. The two dimensional warranty is characterised by a region in two dimensional plane with one axis representing age and the other one usage. Murthy et. al. (1990) deals with two-dimensional renewal processes to model the item failure behaviour under the free replacement and no repair assumption. Chen and Popova(2002) proposes a new maintenance policy which minimizes the total expected servicing cost for an item with two-dimensional warranty.

The majority of the literature on warranty servicing deals with the case where all failures are rectified by corrective maintenance. Also there are different optimal repair-replace strategies and cost limit replacement strategies studied by many researchers. When there is a long warranty period it may be quite worthwhile for the manufacturer to schedule preventive maintenance during the warranty period in order to reduce the chance of system failures. For a brief literature on preventive maintenance one may refer to Jack and Murthy (2002). In the existing literature, optimal maintenance schedules are found by minimizing expected total costs over a finite/infinite time horizon.

Preventive maintenance of product with warranty under general failure time distribution is rather an unexplored area of research. Most devices when used in different places exhibit different levels of deterioration and lifetimes owing to the prevailing environmental conditions which vary in time. This necessitates the use of mixture models for the lifetime of such devices. We consider an attractive model where the underlying life distributions are phase-type and suggest a method of finding the optimal free replacement warranty period with $'n'$ number of preventative maintenances under the given cost restriction , ensuring a specified reliability for the product during the warranty period. By a product we mean either a component or a system throughout the discussion.

In many real life situations, the ageing plays a crucial role with regard to the reliability of a product. When there is a positive ageing the product will have an adverse effect, reducing its reliability. In such situations, by imparting proper maintenances, one may enhance the reliability to a desired level. When the life of the product varies quiet widely, owing to the extraneous factors such as the environmental conditions, place of installation,etc., it is an extremely important and crucial problem faced by the manufacturer to model the life time and also to find the most plausible warranty period for the product. The manufacturer would naturally expect high reliability of the system. The following is a typical situation of a general problem that we are faced with.

Suppose that a manufacturer produces a particular product having random life time. The product seem to have different life time bahaviour over different places where they are used. The manufacturer can amply offer free maintenances (service) while assuring a reliability of at least a pre-decided 'p' percentage. The problem is to obtain the optimal warranty time for the product under a combination of FRW and RI warranty policy so as ensuring a reliability of at least 'p' percentage during the warranty period with $'n'$ free maintenances under the given cost restriction, that the expected cost of warranty should not exceed an allotted amount C. Let X denote the life time of the product under normal (factory) condition. When it is used at a different place the life time becomes $\beta X, 0 < \beta < 1$ where β refers to the degree of reduction in the life time owing to that particular place.

We model the product life times using phase type distributions. The properties (2.5.2),(2.5.6), (2.5.8) of phase type distribution cited in chapter 2, are used for further development in this chapter and we used MATLAB software for the computation of various quantities of interest.

In section 2 we present the problem of obtaining the optimal warranty period for a component, and that for a series system presented in section 3. The method of computing the optimal warranty period is illustrated using a hypothetical data in the last section.

3.2 Determination of warranty period for a component

For the sake of simplicity we shall first consider the case of determining the warranty period of a component manufactured by a company. The manufacturer supplies the components to 'r' different places including the place with factory conditions. Let X be the life time of the component at the factory condition and $X \sim CPH(\underline{\alpha}, T)_m$. Let $Y_1, Y_2, \ldots, Y_{r-1}$ respectively be the life times at (r-1) different places where $Y_i = \beta_i X$, for i=1,2,...,r-1; $0 < \beta_i < 1$. Here $\beta_1, \beta_2, \ldots, \beta_{r-1}$ refer to the factors denoting the effect of the different places on the life of the component. Since X is assumed to have CPH, by the properties of phase type distribution, $Y_i's$ are CPH variates with respective parameters $(\underline{\alpha}, \beta_i^{-1}T)_m$. Let $(\delta_0, \delta_1, \ldots, \delta_{r-1})$ be the sale distribution of the component at the above r places, where δ_0 corresponding to the place of factory, and in general δ_i corresponds to the place with component life Y_i. Now the life time distribution of the component is given by

$$F_Z(t) = P(Z \leq t) = \delta_0 F_X(t) + \sum_{i=1}^{r-1} \delta_i F_{Y_i}(t).$$

It follow from the property that $Z \sim CPH(\underline{\gamma}, L)$, where

$\underline{\gamma} = (\delta_0\underline{\alpha}, \delta_1\underline{\alpha}, \ldots, \delta_{(r-1)}\underline{\alpha})$ and

$$L = \begin{bmatrix} T & 0 & 0 & \ldots & 0 \\ 0 & \beta_1^{-1}T & 0 & \ldots & 0 \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & \beta_{r-1}^{-1}T \end{bmatrix}$$

.

Suppose that the manufacturer decides to offer a FR warranty with $'n'$ free maintenances so as to keep the total warranty cost not exceeding a pre-fixed amount, say $'C'$ and ensuring a reliability of at least $'p'$ percentage during the warranty period.

To compute the warranty period we shall proceed as follows: From the reliability function $R(t)$ of the lifetime Z following CPH$(\underline{\gamma}, L)$ where,

$$R(t) = P(Z > t) = \underline{\gamma}e^{Lx}\underline{e},$$

we find a time point t_1, such that $t_1 = Sup\{t/R(t) \geq p\}$. At this time point t_1 the manufacturer provide the first maintenance service to the component which boosts its life. But because of ageing it cannot act as a new component. The maintenance time is small relative to the mean working time of the component and so can be ignored. After the first maintenance, the life time of the the component at the place of factory condition is denoted by X_1, where $X_1 = q_1X$, and the life at the other $r-1$ different places as Y_{i1}, where $Y_{i1} = q_1Y_i$; i=1,2,..., (r-1). where q_1 represent the joint effect of ageing and maintenance on the life time of the component after the first maintenance. In general it is represented as q_j for the component after the j^{th} maintenance. The lifetimes of the components after the j^{th} maintenance at the place of factory condition and at the remaining $(r-1)$ places are respectively denoted by X_j and Y_{ij}, for $i = 1, 2, \ldots, (r-1); j = 1, 2, \ldots$,n. As in the last case, we can find the probability distribution of the life time Z_1 of the component after the first maintenance as a mixture by,

$$F_{Z_1}(t) = P(Z_1 \leq t) = \delta_0 F_{X_1}(t) + \sum_{i=1}^{r-1} \delta_i F_{Y_{i1}}(t).$$

Here $X_1 \sim CPH(\underline{\alpha}, q_1^{-1}T)$ and $Y_{i1} \sim CPH(\underline{\alpha}, (q_1\beta_i)^{-1}T)$ for i= 1,2,...,(r-1). So we have $Z_1 \sim CPH(\underline{\gamma}, L_1)$ where

$$L_1 = \begin{bmatrix} q_1^{-1}T & 0 & 0 \ldots & 0 \\ 0 & (q_1\beta_1)^{-1}T & 0 \ldots & 0 \\ . & . & . \ldots & . \\ . & . & . \ldots & . \\ 0 & 0 & 0 \ldots & (q_1\beta_{r-1})^{-1}T \end{bmatrix}.$$

Using the life distribution of the component after the first maintenance, plot its reliability function and find the next time point t_2 for giving the second free maintenance, where $t_2 = Sup\{t/R_1(t) \geq p\}, R_1(t)$ being the reliability function associated with the life time Z_1. As before we can find the life time of the component, after the second maintenance by considering the joint effect of the ageing and maintenance on the life time of the component as q_2. Thus we can derive the probability distribution of the lifetime Z_2 of the component after the second maintenance as $CPH(\underline{\gamma}, L_2)$where,

$$L_2 = \begin{bmatrix} q_2^{-1}T & 0 & 0 & \ldots & 0 \\ 0 & (q_2\beta_1)^{-1}T & 0 & \ldots & 0 \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & (q_2\beta_{r-1})^{-1}T \end{bmatrix}$$

Next we find the time point t_3, where $t_3 = Sup\{t/R_2(t) \geq p\}$; $R_2(t)$ is the reliability function of the component after the second maintenance. The component works up to this time point t_3 with atleast p percentage of reliability and this is the time for the third maintenance. Continuing like this one can find the time points for the successive maintenances so as to ensure the component is working with atleast p percentage of reliability. Then the component is considered working upto a time $t_{(n+1)}$ with $'p'$ percentage of reliability after the n^{th} maintenance. Hence the total time the component is expected to work with at least 'p' percentage of the reliability is $t = t_1 + t_2 + \ldots + t_{(n+1)}$. The manufacturer restricts that the total warranty cost with maintenances not to exceed a pre fixed amount $'C'$. So it is to find the optimal number of maintenances and optimal warranty period within the restriction that the expected warranty cost $E(C_w) \leq C$. Expected cost of warranty with $'n'$ maintenances can be expressed as

$$E(C_w) = C_s(M_1(t_1) + M_2(t_2) + \ldots + M_{(t_{n+1}(t_{n+1}))}) + nC_m \tag{3.2.1}$$

where, C_s is the component cost, $M_i(t_i)$ is the expected number of failures in the duration t_i after the $(i-1)^{th}$ maintenance. C_m is the cost per maintenance which is assumed uniform for all of the maintenances. Expected number of failures up to time t of a component can be expressed as

$$M(t) = \int_0^t h(t)dt = -log\overline{F}(t). \tag{3.2.2}$$

where, h(t) is the hazard function and F is the life distribution of the component. The expected number of failures between each of the maintenances represented by $M_1(t_1), M_2(t_2), \ldots, M_{(t_{n+1})}$ can be found using the corresponding estimated CPH life distribution. The restriction $E(C_w) \leq C$, is verified for $n = 1, 2, \ldots$. The highest value of n satisfying the restriction $E(C_w) \leq C$ is the optimal number of maintenances and with respect to the total number of maintenances $'n'$, the optimal warranty period that can offer is $t = t_1 + t_2 + \ldots + t_{(n+1)}$.

3.3 Warranty period for a series system

The foregoing methods can be adopted for optimizing the warranty period of a system with the components arranged serially. Let there be $'k'$ components for the system with lifetime $X_1, X_2 \ldots\ X_k$ as continuous phase type(CPH) distributed with respective parameters $(\underline{\alpha}_1, T_1)_{m_1}, (\underline{\alpha}_2, T_2)_{m_2}, \ldots\ (\underline{\alpha}_k, T_k)_{m_k}$ in the factory condition or in the standard situation. Suppose that this company supplies the system to r different places, including the place with standard situation. Out of the r different places, let the first place provide standard situation for the system where the system life is denoted by X_{00} and the system life at the remaining $(r-1)$ places are denoted by Y_{i0}, for i = 1,2, . . . (r-1). At the first place other than the place with factory condition, let the life of the components of the system denoted by $\beta_{11}X_1, \beta_{12}X_2 \ldots \beta_{1k}X_k$., where β_{1j} is the effect on the life of the j^{th} component due to the physical atmosphere of the second place, where $0 < \beta_{1j} < 1$, j= 1,2,.....k. Then the life of the components in this situation follows CPH with parameters $(\underline{\alpha}_1, \beta_{11}^{-1}T_1)_{m_1}, (\underline{\alpha}_2, \beta_{12}^{-1}T_2)_{m_2}, \ldots\ (\underline{\alpha}_k, \beta_{1k}^{-1}T_k)_{m_k}$ respectively.

In general, let the life of the j^{th} component working at i^{th} place other than the place with standard condition, is denoted as $\beta_{ij}X_j$ where $0 < \beta_{ij} < 1$, i=1,2,.....(r-1) ; j=1,2,.....k; which is the effect on the life of the j^{th} component working at the i^{th} place. At the place where $i = (r-1)$, the life of the components are denoted as $\beta_{(r-1)1}X_1, \beta_{(r-1)2}X_2 \ldots \beta_{(r-1)k}X_k$.,. Consequently the life of the k components at this place follows Continous phase type distribution with paraments

$(\underline{\alpha_1}, \beta^{-1}_{(r-1)1}T_1)_{m_1}, (\underline{\alpha_2}, \beta^{-1}_{(r-1)2}T_2)_{m_2}, \ldots (\underline{\alpha_k}, \beta^{-1}_{(r-1)k}T_k)_{m_k}$, respectively.

Now for series system under consideration, the life time in the standard environment or in factory condition X_{00} is the minimum of the life times of the components which follows CPH with respective parameters $(\underline{\alpha_1}, T_1)_{m_1}, (\underline{\alpha_2}, T_2)_{m_2}, \ldots\ldots$ $(\underline{\alpha_k}, T_k)_{m_k}$, which is again a CPH with parameters $(\underline{\alpha_{00}}, T_{11})_{m_1 \times m_2 \ldots \times m_k}$, where

$$\underline{\alpha_{00}} = [\alpha_1 \otimes \alpha_2 \otimes \alpha_3 \ldots \otimes \alpha_k], and$$

$$T_{00} = [T_1 \otimes I_{m_2 \times m_3 \times \ldots \times m_k} + I_{m_1} \otimes T_2 \otimes I_{m_3 \times m_4 \times \ldots \times m_k} + \ldots + I_{m_1 \times m_2 \times \ldots \times m_{k-1}} \otimes T_k]$$

where I_{m_1} is an identity matrix of order m_1. Similarly, the distribution of the life of the system at the first place, other than the place with standard condition Y_{10} is a CPH variate with parameters $(\underline{\alpha_{00}}, T_{10})_{m_1 \times m_2 \ldots \times m_k}$. where,

$$T_{10} = [\beta_{11}^{-1}T_1 \otimes I_{m_2 \times m_3 \times \ldots \times m_k} + I_{m_1} \otimes \beta_{12}^{-1}T_2 \otimes I_{m_3 \times m_4 \times \ldots \times m_k} + \ldots + I_{m_1 \times m_2 \times \ldots \times m_{k-1}} \otimes \beta_{1k}^{-1}T_k]$$

Continuing in this way we get the life of the system at the $(r-1)^{th}$ place $Y_{(r-1)0}$ as a phase type variate with parameters$(\underline{\alpha_{00}}, T_{(r-1)0})$, where

$$T_{(r-1)0} = [\beta^{-1}_{(r-1)1}T_1 \otimes I_{m_2 \times m_3 \times \ldots \times m_k} + I_{m_1} \otimes \beta^{-1}_{(r-1)2}T_2 \otimes I_{m_3 \times m_4 \times \ldots \times m_k} + \ldots$$
$$+ I_{m_1 \times m_2 \times \ldots \times m_{k-1}} \otimes \beta^{-1}_{(r-1)k}T_k].$$

Suppose that systems produced by the factory are sent to the r places randomly with respective probablities $\delta_0, \delta_1, \ldots, \delta_{(r-1)}$. Since the system has to work under any one of these randomly selected places, distribution function of life Z of a system, can be taken as a mixture of the distribution functions of the lifes of the system under various places with the mixing density $\delta_0, \delta_1, \ldots, \delta_{(r-1)}$. We know at the i^{th} place, the distribution function of the system life

$F_{Y_{i0}}(t) = 1 - \underline{\alpha}_{00} e^{T_{i0}t} \underline{e}$,where e=(1,1,...,1). Consequently the life time of the system Z has the distribution function,

$$F_Z(t) = \delta_0 F_{X_{00}}(t) + \sum_{i=1}^{r-1} \delta_i F_{Y_{i0}}(t), \tag{3.3.1}$$

where $Y_{i0} \sim CPH(\underline{\alpha}_{00}, T_{i0})$. But a finite mixture of CPH variates is again a CPH variate. Note that $Z \sim CPH$ distribution with parameters (γ, L)

where $\gamma = (\delta_0 \alpha_{00}, \delta_1 \alpha_{00}, \ldots, \delta_{r-1} \alpha_{00})$ and

$$L = \begin{bmatrix} T_{00} & 0 & 0 & \ldots & 0 \\ 0 & T_{10} & 0 & \ldots & 0 \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & T_{(r-1)0} \end{bmatrix}$$

Now we can obtain the reliability of the system and its hazard rate at different time point.

Here we want to find how far the system can run with a reliability atleast $'p'$ percentage. The system starts at time t=0, and we can find the highest value of t, say t_1 such that $R(t) \geq p$, where $R(t)$ is the reliability of the system with the life time Z at time t. But here the manufacturer decided to give an extended waranty period by giving maintenance services. So at time t_1 of running of the system, the manufacturer should provide the first maintenance to all of the components of the system uniformly. This maintenance boosts the efficiency of the components, but even though a maintenance is given it is natural that it cannot work as new due to ageing. The life times of the system after the j^{th} maintenance at the factory condition is denoted by X_{0j} and at the remaining (r-1) places are denoted by Y_{ij} for $i = 1, 2, \ldots, (r-1), \quad j = 1, 2, \ldots$. After the first maintenance, the life length of a components under standard condition is distributed as continuous phase type with respective parameters $(\underline{\alpha}_1, q_{11}^{-1}T_1)_{m_1}, (\underline{\alpha}_2, q_{12}^{-1}T_2)_{m_2}, \ldots \ (\underline{\alpha}_k, q_{1k}^{-1}T_k)_{m_k}$ where q_{1i} is the joint effect of the ageing and maintenance on the life of the i^{th} component after the first maintenance. $(0 < q_{1i} < 1$, for i = 1,2,.......k) and hence, the life of the system after the maintenance at the factory condition,X_{01}, is the minimum of the component life times and follows CPH with parameters $(\underline{\alpha}_{00}, T_{01})$ where,

$$T_{01} = [q_{11}^{-1}T_1 \otimes I_{m_2 \times m_3 \times \ldots \times m_k} + I_{m_1} \otimes q_{12}^{-1}T_2 \otimes I_{m_3 \times m_4 \times \ldots \times m_k} + \ldots + I_{m_1 \times m_2 \times \ldots \times m_{k-1}} \otimes q_{1k}^{-1}T_k]$$

In accordance with this, similar effect in life distribution after the first maintenance is happening to all the components in running at all of the different places. So at the i^{th} destination other than the place with factory condition, the components life after the first maintenance may be distributed as CPH with respective parameters $(\underline{\alpha}_1, (q_{11}\beta_{i1})^{-1}T_1)_{m_1}, (\underline{\alpha}_2, (q_{12}\beta_{i2})^{-1}T_2)_{m_2}, \ldots \ (\underline{\alpha}_k, (q_{1k}\beta_{ik})^{-1}T_k)_{m_k}$. So the life of the system after the first maintenance at the i^{th} place Y_{i1} is the minimum of the life of these components, which is again CPH with parameters$(\underline{\alpha}_{00}, T_{i1})_{m_1 \times m_2 \ldots \times m_k}$.

where,

$$T_{i1} = [(q_{11}\beta_{i1})^{-1}T_1 \otimes I_{m_2 \times m_3 \times \ldots \times m_k} + I_{m_1} \otimes (q_{12}\beta_{i2})^{-1}T_2 \otimes I_{m_3 \times m_4 \times \ldots \times m_k}$$

$$+ \ldots + I_{m_1 \times m_2 \times \ldots \times m_{k-1}} \otimes (q_{1k}\beta_{ik})^{-1}T_k],$$

for i=1,2,...,(r-1).

Therefore as shown above, we can find the distribution of the life of the system after the first maintenance as a mixture of life distributions of different places which again a CPH variate (let it be Z_1) with parameters (γ, L_1),where γ is as given above, but L_1 is

$$L_1 = \begin{bmatrix} T_{01} & 0 & 0 & \ldots & 0 \\ 0 & T_{11} & 0 & \ldots & 0 \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & T_{(r-1)1} \end{bmatrix}$$

By using the distribution of Z_1 we can find the reliability of the system after the first maintenance. Hence again we can find a time point, t_2 such that $t_2 = Sup\{t/R_1(t) \geq p\}$. Here $R_1(t)$ is the reliability function of the system with life time Z_1. This is the time of running of the system from the first maintenance,where the manufacturer have to provide the second maintenance. Considering the joint effect of ageing and maintenance on the life of i^{th} component after the second maintenance as q_{2i}. This changes the life time of a component under standard condition and is distributed as continuous phase type with respective parameters $(\underline{\alpha_1}, q_{21}^{-1}T_1)_{m_1}, (\underline{\alpha_2}, q_{22}^{-1}T_2)_{m_2}, \ldots \ (\underline{\alpha_k}, q_{2k}^{-1}T_k)_{m_k}$. Then as we proceed in the case of the life distribution after the first maintenance we can derive the distribution of the life of the system after the second maintenance Z_2 as a phase type variate with parameters (γ, L_2), where,

$$L_2 = \begin{bmatrix} T_{02} & 0 & 0 & \ldots & 0 \\ 0 & T_{12} & 0 & \ldots & 0 \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & T_{(r-1)2} \end{bmatrix}$$

$$T_{02} = [q_{21}^{-1}T_1 \otimes I_{m_2 \times m_3 \times \ldots \times m_k} + I_{m_1} \otimes q_{22}^{-1}T_2 \otimes I_{m_3 \times m_4 \times \ldots \times m_k} + \ldots\ldots + I_{m_1 \times m_2 \times \ldots \times m_{k-1}} \otimes q_{2k}^{-1}T_k]$$

and for i= 1,2,....,(r-1),

$$T_{i2} = [(q_{21}\beta_{i1})^{-1}T_1 \otimes I_{m_2 \times m_3 \times \ldots \times m_k} + I_{m_1} \otimes (q_{22}\beta_{i2})^{-1}T_2 \otimes I_{m_3 \times m_4 \times \ldots \times m_k} + \ldots + I_{m_1 \times m_2 \times \ldots \times m_{k-1}} \otimes (q_{2k}\beta_{ik})^{-1}T_k]$$

From the graph of the reliability function of the distribution of Z_2 we can find a time point, t_3 such that $t_3 = Sup\{t/R_2(t) \geq p\}$. Here $R_2(t)$ is the reliability function of the system with life time Z_2. Similarly one can find the time points for the $'n'$ maintenances as $t_4, t_5, \ldots, t_n$. Making use of the estimated life distribution of the system in different time intervals and using (3.2.1) one can find the optimal number of maintenances to be provided within the cost restriction and the optimal warranty can be found as described earlier in the case of component.

Remark 3.3.1 *The procedure outlined above can be implemented for the case of a parallel system or any general coherent system in order to find out the optimal warranty period.*

3.4 Numerical Illustration

Consider a hypothetical situation where a component with lifetime X following $CPH(\underline{\alpha}, T)$ with $\underline{\alpha} = (1,0,0)$ and

$$T = \begin{bmatrix} -.6 & .1 & .2 \\ .5 & -1.7 & .1 \\ .1 & .4 & -1.6 \end{bmatrix}$$

under the factory condition, being used at three different places with the parameter values $\beta_1 = 0.8$, $\beta_2 = 0.5$. The desired reliability level while offering a warranty of atleast 80% and the component sale distribution among the three different places is assumed to be {0.7,0.1,0.2} where the first entry corresponds to the one having factory condition. The cost of the component is assumed to be 10,000, C_m=150 and the warranty cost limit is C=5,000.

The foregoing method yields the life of the system as following $CPH(\gamma, L)$, where $\gamma = (1, 0, 0, 1, 0, 0, 1, 0, 0)$ and

$$L = \begin{bmatrix} T & 0 & 0 \\ 0 & T/0.8 & 0 \\ 0 & 0 & T/0.5 \end{bmatrix}$$

The plot of the reliability function (figure 1) using MATLAB for this life distribution, gives $t_1 = 0.5$. Using (3.2.2), the expected number of failures up to t_1 is calculated as $M_1(t_1) = 0.2177$. Hence for $r = 1$, $E(C_w) = 2327$ which is not greater that the given value of C. So the first maintenance is given at the time point t_1 and the life of the system after the maintenance for the value of $q_1 = 0.8$ is following $CPH(\gamma, L_1)$ where,

$$L_1 = \begin{bmatrix} T/0.8 & 0 & 0 \\ 0 & T/(0.8 \times 0.8) & 0 \\ 0 & 0 & T/(0.8 \times 0.5) \end{bmatrix}$$

The reliability function of this life in figure 2 gives the time upto where the component to work with atleast 80 percentage of reliability after this maintenance as $t_2 = 0.4$. The expected number of failures in this time span is calculated as $M_2(t_2) = 0.2178$ and hence for $r = 2$, $E(C_w)$ is calculated as 4657 which is less than C. Then is is allowed to provide the second maintenance at this time point t_2. The life of the component after this maintenance for the value of $q_2 = 0.6$. is

$CPH(\gamma, L_2)$ where,

$$L_2 = \begin{bmatrix} T/0.6 & 0 & 0 \\ 0 & T/(0.6 \times 0.8) & 0 \\ 0 & 0 & T/(0.6 \times 0.5) \end{bmatrix}$$

The plot of reliability function for the product after this maintenance in figure 3 gives t_3=0.34, that is the component works up to this time point after the second maintenance with the required reliability and the expected number of failures in this region is calculated as 0.2502. The possibility of the next maintenance is verified by $E(C_w)$ for $r = 3$. On calculation $E(C_w)$ =7306, which is greater that the given value of C. Therefore only two maintenances are possible and with these two maintenances the manufacturer can offer a warranty for the period $W = t_1 + t_2 + t_3 = 1.24$ units of time.

3.5 Figures:

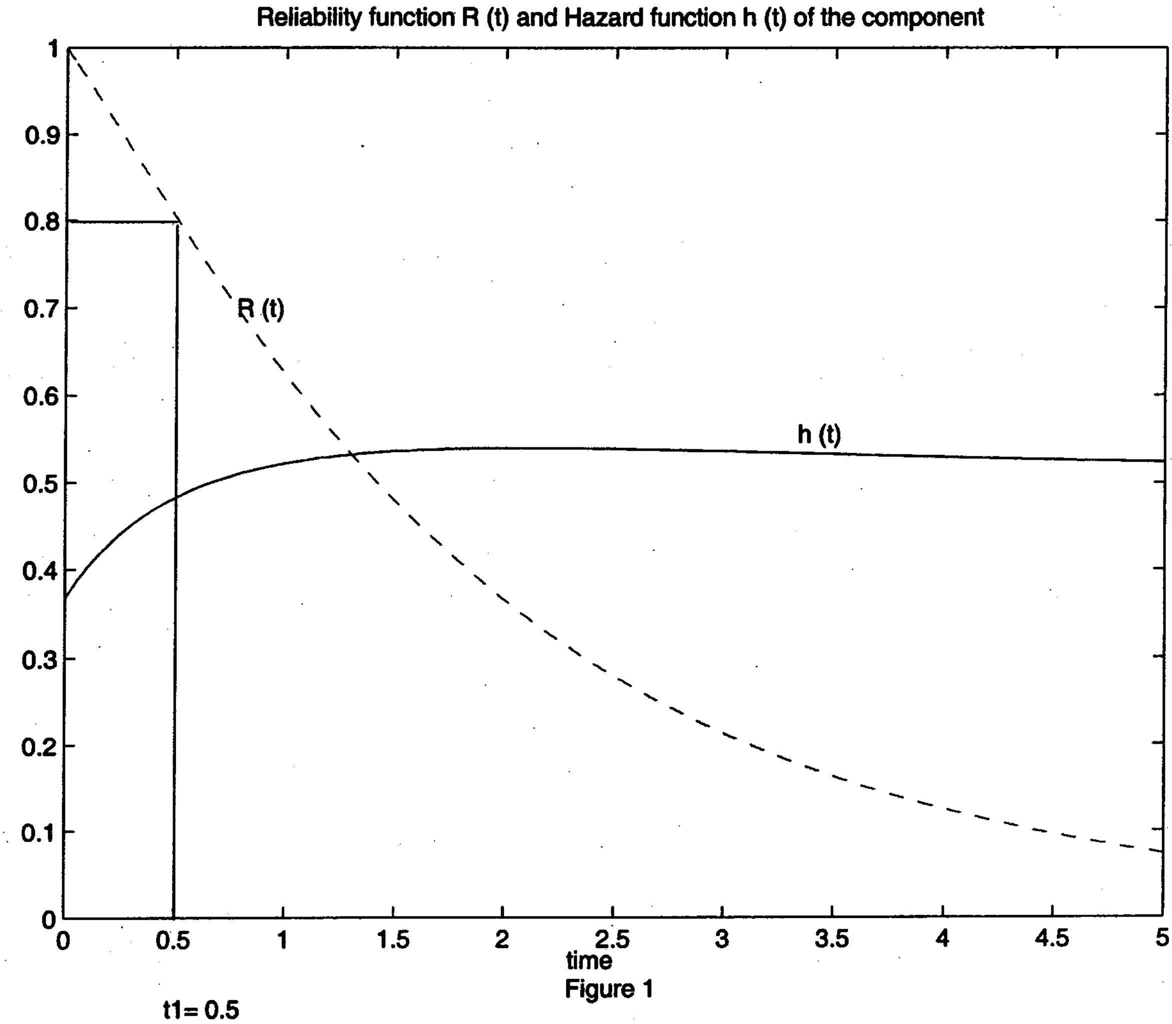
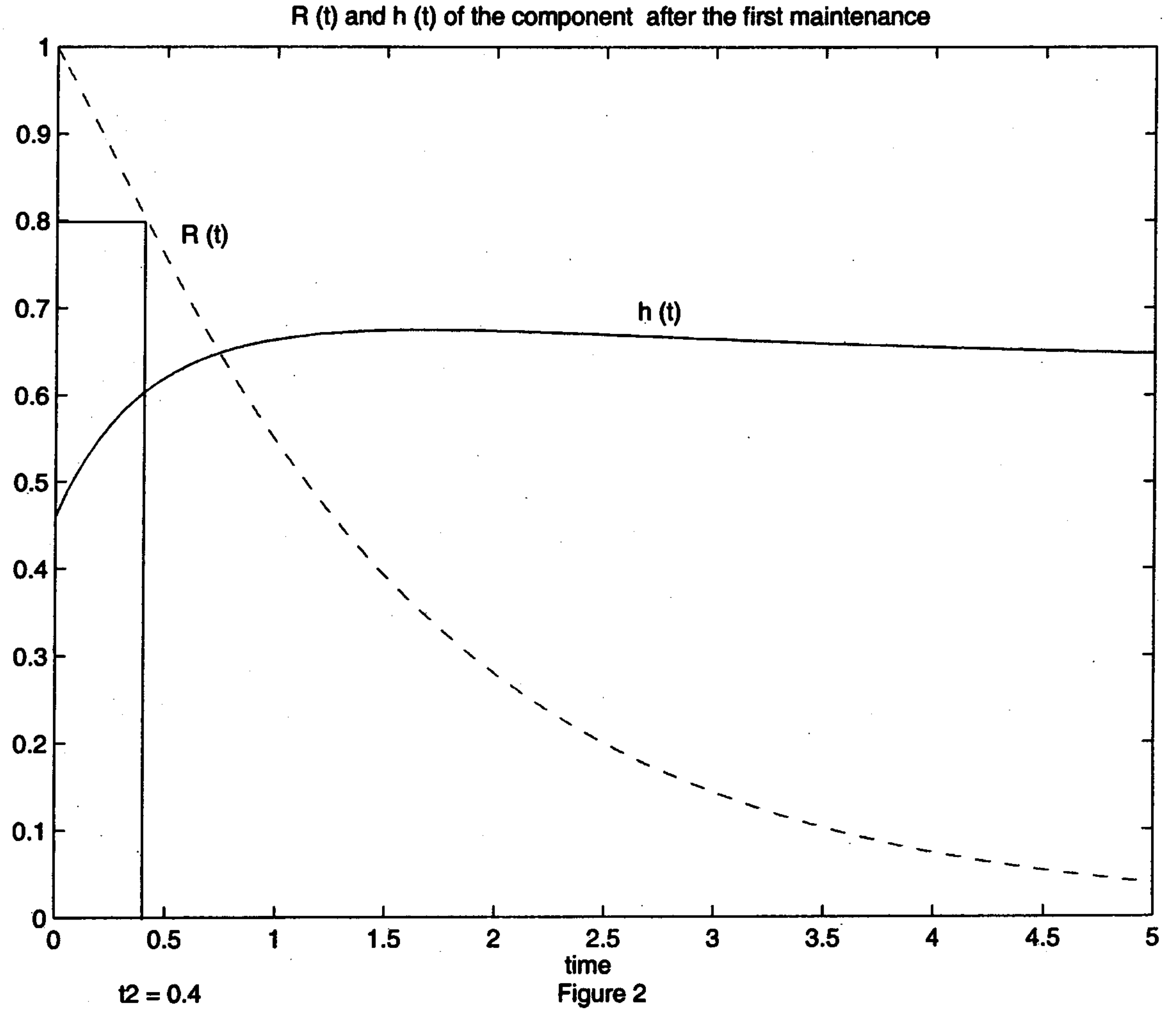
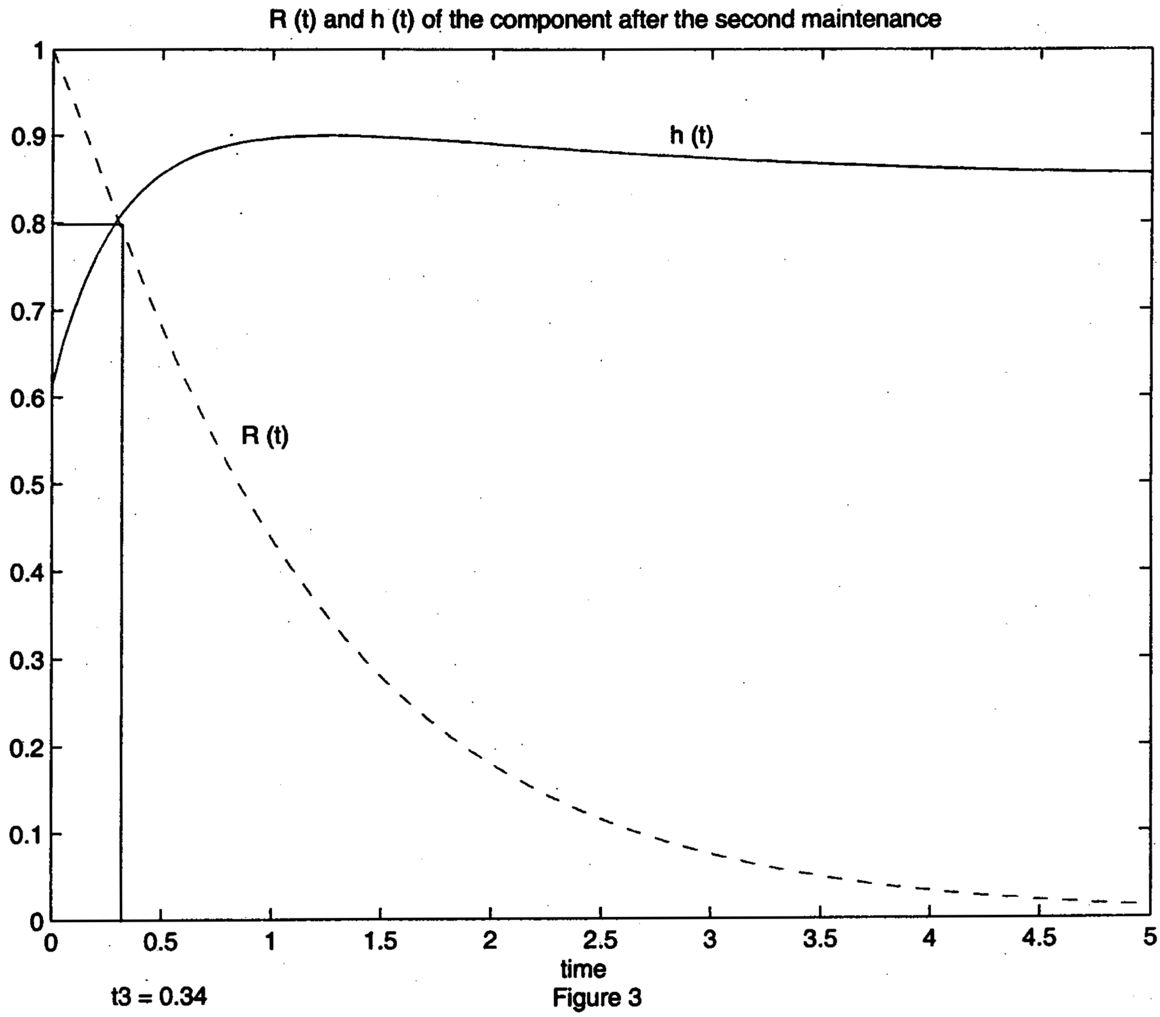


Figure 1


R (t) and h (t) of the component after the first maintenance
R (t)
h (t)
time
t2 = 0.4
Figure 2


R (t) and h (t) of the component after the second maintenance
h (t)
R (t)
time
t3 = 0.34
Figure 3

ON PREVENTIVE REPLACEMENT OF A DEVICE UNDER POISSON SHOCKS WITH DPH SHOCK PROBABILITIES

Aneesh Kumar K. “Phase type modelling in reliability and survival analysis”
Thesis. Department of Ststistics , University of Calicut, 2007

Chapter 4

ON PREVENTIVE REPLACEMENT OF A DEVICE UNDER POISSON SHOCKS WITH DPH SHOCK PROBABILITIES

4.1 Introduction

The optimal replacement problem has been the centre of attraction to many researchers for a long time,wherein we consider a system that will deteriorate and thus should be replaced by a new one when it is reasonably bad. For example see Cho and Parlar(1991), Manoharan and Joseph(2004). There are two types of the deterioration considered in the reliability literature, one is the continuous time deterioration due to the operation of the system while the other is caused by influence of the environment; for example, shocks to the system. These are known as system deterioration and the environment deterioration respectively. In general, the degradation process of a system can be controlled by monitoring or by inspection - continuous or periodic. Maintenance action can be applied when deterioration exceeds a particular threshold as described by Berenguer et. al.(2003)

However there are many situations where it is not possible to examine or inspect the current state of the device. In such cases a policy for preventive replacement of the device is highly desirable. That is to find a pre-fixed replacement age for the device so as to minimize the long run cost per unit time. Several studies have been directed towards this problem in the past. Recently Suprasad Amari and Wes Fulton(2003) presented bounds on optimal replacement time of age replacement policy. In this thesis we propose a technique to obtain the replacement age of a device under homogenous Poisson shocks with DPH shock probabilities, by minimizing the long run cost per unit time for replacement, and also in another case by considering a particular loss function . We assume that the loss due to preventive replacement of the device is small in comparison with the loss due to the corrective replacement (replacement after failure).

Consider a device subject to shocks governed by a homogenous Poisson process. Let p_k be the probability that the device fails at the k^{th} shock and,

$$\overline{P_k} = \sum_{n=k}^{\infty} p_k$$

be the probability that the device survive the k^{th} shock, where $1 = \overline{P_0} \geq \overline{P_1} \geq \overline{P_2} \geq \ldots$ Now if $\{p_k\}$ have been specified, it is relatively easy to find the life distribution of the system. Manoharan et. al.(1992) have noticed that if $\{p_k\}$ has a discrete phase type distribution, then the life distribution $H(t)$ of the device, subject to shocks governed by a finite mixture of homogeneous Poisson processes, has a continuous phase type distribution. The structural property of this life distribution can be further exploited to gather better performance measures about the device. In particular the optimal replacement age can be obtained using the phase type formalism for the life distribution of the device.

4.1.1 Life of a device under Poisson shocks

The survival function of the device subject to shocks occurring randomly over time according to a counting process $N = \{N(t), t \geq 0\}$ and with the probability of survival of k shocks $\overline{P_k}$ is

$$\overline{H}(t) = \sum_{k=0}^{\infty} P\{N(t) = k\}\overline{P_k}. \tag{4.1.1}$$

Shock models of this kind have been extensively studied in the past and it has been verified that the distribution function of the lifetime of the device $H(t)$ has IFR,IFRA,DMRL,NBU,NBUE and HNBUE property if the shock probabilities $\{p_k\}$ has the corresponding discrete property. Note that when device is under homogenous Poisson shocks with parameter λ, the above expression takes the form,

$$\overline{H}(t) = \sum_{k=0}^{\infty}[\frac{e^{-\lambda t}[\lambda t]^k}{k!}]\overline{P_k}. \tag{4.1.2}$$

Now we study the survival function of the device under the shock probabilities $\{p_k\}$ follows a discrete phase type distribution (DPH), a versatile class of distributions on the positive integers.

4.1.2 Poisson shocks with Discrete Phase type shock probabilities

There are several real life situations where one never knows the precise form of distribution for the shock probabilities $\{p_k\}$. So in such general situation we can make use of the phase type modelling to solve the problem. The class of phase type distribution is dense, and hence any distribution with positive support can be approximated arbitrarily close by a phase type distribution. [See

Neuts(1981)]. Let us consider a device under homogenous Poisson shocks and the shock probabilities follow discrete phase type distribution with parameters (α, T). In such a case Manoharan et.al.(1992) established that the survival function of the device can be represented as

$$\overline{H}(t) = \alpha e^{\lambda(T-I)t} \underline{e}, \tag{4.1.3}$$

where λ is the rate of Poisson shocks. So the life distribution of the device can be represented as a continuous phase type distribution(CPH) with parameters α and $\lambda(T-I)$.

Special case:

Consider a device, which is under homogenous Poisson shocks with parameterλ. Suppose that the shock probability distribution $\{p_k\}$ is a geometric distribution-which is a particular case of discrete phase type distribution when the parameters are $\alpha = (1)$ and $T = (q)$. Then $p_k = q^{k-1}p$, 'q' being the probability to survive a shock and while 'p' the probability of failure by a shock, which are assumed same for all shocks and shock effects are considered independent. Here the probability of the system to survive k^{th} shock is

$$\overline{P_k} = \sum_{k=k+1}^{\infty} q^{k-1}p = q^k.$$

Therefore the survival function of the device is,

$$\overline{H(t)} = \sum_{k=o}^{\infty}[\frac{e^{-\lambda t}[\lambda t]^k}{k!}]q^k$$

$$= e^{-\lambda t(1-q)} \qquad (4.1.4)$$

which is in the form of the survival function of exponential distribution with parameter $\lambda(1-q)$. Hence the life of the device follow exponential distribution which is the continuous phase type life with parameters $\alpha = (1), T = -\lambda(1-q)$.

4.2 Optimal Replacement Age of the Device

In this chapter we present two different situations of finding optimal replacement age for a device under Poisson shock with DPH shock probabilities. To find the optimal replacement age of such a device we use the result that the lifetime of the device is CPH distributed. The first method suggested is that based on cost aspects, in which the optimal age of the device is obtained by minimizing the long run cost of replacement per unit time. The second method assumes the prior knowledge of the loss function in a replacement situation. That is, we consider a situation of age replacement, where the device under Poisson shock is decided to replace before its expected lifetime for minimizing the risk due to failure at its maximum. In this case, the possible loss due to an early replacement of the device is formulated. The loss function $L(t)$ denotes the possible loss due to replacement at time 't'. The optimum replacement age is the time 't' which minimizes $L(t)$. In both the cases the gain due to age replacement is quantified.

4.2.1 Optimal replacement age by minimising long run cost per unit time for replacement

Consider the device under homogenous Poisson shocks with parameter λ and with shock probabilities follows discrete phase type distribution with parameter (α, T). The life of the device is derived as following continous phase type distributions with parameters $(\alpha, \lambda(T - I))$. The device is replaced either upon failure or when its age reaches some fixed time t, which ever comes first. When replacement occurs, the new unit is identical to the replaced unit. The replacement cost before failure is C_1 and the cost of replacement due to failure is C_2, where $C_1 < C_2$. In such a situation, the long run cost per unit time for replacement at time t, of a device with life $F(.)$ is

$$C_a = \frac{C_1(1 - F(t)) + C_2(F(t))}{\int_0^t (1 - F(x))dx} \qquad (4.2.1)$$

To find the optimal age t for replacement we minimise C_a with respect to $'t'$. Applying differential calculus for finding optimal t, we have,

$$\int_0^t (1 - F(x))dxf(t)(C_2 - C_1) = C_1(1 - F(t)) + C_2F(t)(1 - F(t))$$

$\Rightarrow$

$$\frac{f(t)}{1 - F(t)} \int_0^t (1 - F(x))dx = \frac{C_1}{C_2 - C_1} + F(t) \qquad (4.2.2)$$

$\Rightarrow$

$$h(t)\int_0^t (1-F(x)dx - F(t) = \frac{C_1}{C_2 - C_1} \tag{4.2.3}$$

where $h(t)$ denotes the hazard rate. Now using the fact that the lifetime of the device follows $CPH(\alpha, \lambda(T-I))$, we have

$$\begin{aligned} F(t) =& 1 - \alpha e^{\lambda(T-I)t}\underline{e} \\ h(t) =& \frac{\alpha e^{\lambda(T-I)t}T_0}{\alpha e^{\lambda(T-I)t}\underline{e}}. \end{aligned}$$

Then equation (3.7) becomes,

$$\frac{\alpha e^{\lambda(T-I)t}T_0}{\alpha e^{\lambda(T-I)t}\underline{e}}\int_0^t \alpha e^{\lambda(T-I)x}\underline{e}dx - [1-\alpha e^{\lambda(T-I)t}\underline{e}] = \frac{C_1}{C_2 - C_1} \tag{4.2.4}$$

The value of t found by solving equation (4.2.4) is suggested as the optimal replacement time of the device.

Numerical Illustration

Consider a device under homogenous Poisson shock with parameter $\lambda = 2$, and with the shock probabilities following $DPH(\alpha, T)$, where

$$\alpha = (1, 0, 0)$$

$$T = \begin{bmatrix} 0.75 & 0.15 & 0.05 \\ 0.15 & 0.65 & 0.20 \\ 0.30 & 0.10 & 0.55 \end{bmatrix}$$

Then it follows that the life of the device has $CPH(\alpha, T)$, where

$$\alpha = (1,0,0)$$

$$T = \begin{bmatrix} -0.5 & 0.3 & 0.1 \\ 0.3 & -0.7 & 0.2 \\ 0.6 & 0.2 & -0.9 \end{bmatrix}.$$

It is verified that the distribution considered is IFR. Now for a trial calculation, we take $C_1 = 10$ and $C_2 = 100$. The MATLAB program gives us the value of t satisfying the equation (4.2.4) as a value lying in the interval [0.1242,0.1243]. Hence a three decimal approximation of the replacement time t for the device can be taken as 0.124 units of time.

The gain due to age replacement

In this section we compare the preventive replacement policy considered in this paper with the failure replacement policy in terms of the expected long run cost per unit time. Assume one follows the age replacement policy for replacing a device at the time 't'. Let Y_k denote the cost of kth replacement. Let C_1 be the cost of preventive replacement (replacement prior to failure) at 't' and C_2 be the cost of replacement in case of corrective replacement (replacement after failure). This implies

$$Y_k = \begin{cases} C_1 & \text{with probability 1 - F(t),} \\ C_2 & \text{with probability F(t).} \end{cases}$$

Therefore the long run cost per unit time in this case of age replacement is,

$$C_a = \frac{C_1(1-F(t)) + C_2(F(t))}{\int_0^t (1-F(t))dt} \tag{4.2.5}$$

If one does not follow the age replacement policy, but decides to replace the device with a new similar one, only after the device fails, then for a finite time period 't', there may occur a finite number of renewals of the device. The expected long run cost per unit time in this case can be estimated as

$$C_f = \frac{C_2 \times \text{Expected number of renewals in (0,t)}}{t} \tag{4.2.6}$$

Assume that the life distribution of the device is $CPH(\alpha, T)$. Then the associated renewal process is a phase type renewal process. Following Neuts (1981), and assuming $\alpha_{m+1} = 0$, we have the expected number of renewals, $N(t)$ as

$$\begin{aligned} E[N(t)] &= \mu_1'^{-1}t - (1-\alpha_{m+1})^{-1} + \frac{\sigma^2 + \mu_1'^2}{2\mu_1'^2} \\ &\quad + \mu_1'^{-1}(1-\alpha_{m+1})^{-1})\alpha[\Pi - exp(Q^* t)]T^{-1}\underline{e} \end{aligned} \tag{4.2.7}$$

where,

μ_n' is the n^{th} raw moment, σ^2 is the variance of $F(.)$, and $Q^* = (T + T^0\alpha)$.

Further, $\mu_n' = (-1)^n n! \alpha T^{-n}\underline{e}$. Here the matrix Π is defined as $\Pi_{ij} = \pi_j$ for all i, where stationary probability vector π_j are the components of $\pi = \mu_1'^{-1}\alpha(-T)^{-1}$.

Consequently the expected number of renewals in (0,t) becomes,

$$E[N(t)] = (-\alpha T^{-1}\underline{e})^{-1}t - 1 + \frac{\alpha T^{-2}\underline{e}}{(\alpha T^{-1}\underline{e})^2} \tag{4.2.8}$$

$$(-\alpha T^{-1}\underline{e})^{-1}\alpha[\Pi - exp(T + T^0\alpha)t]T^{-1}Q^*t)]T^{-1}\underline{e}$$

The gain due to age replacement can be represented as $G = C_f - C_a$, which can be used as a relative measure of performance evaluation of the two replacement policies.

For illustration purpose we consider the example of section 2, wherein we have the life distribution of the device as $CPH(\alpha, T)$ where,

$$\alpha = (1, 0, 0)$$

$$T = \begin{bmatrix} -0.5 & 0.3 & 0.1 \\ 0 & -0.9 & 0.8 \\ 0.2 & 0.3 & -0.6 \end{bmatrix}$$

with the optimal replacement age 0.124 time units, for C_1=10 and C_2=100, we get $C_a = 90.3130$ and $C_f = 281.2802$, yielding the gain, G=190.9672. The following table gives the optimal replacement time t and gain G, for different values of C_1 and C_2

C_1	C_2	t	C_a	C_f	$G = C_f - C_a$
5	25	0.283	20.0003	70.3201	50.3168
15	125	0.153	110.0425	351.6003	241.5578
10	100	0.124	90.3130	281.2802	190.9672
20	125	0.215	104.8634	351.6003	246.7369
25	150	0.225	125.2851	421.9203	296.6352

The numerical results show significant gain due to age replacement in all cases and thereby confirm its superiority.

4.2.2 Optimal replacement age by considering a loss function

For a device, under homogenous Poisson shocks with parameterλ. and the failure probability p_k follows Geometric distribution, where $p_k = q^{k-1}p$, 'q' being the probability to survive a shock and while 'p' the probability of failure by a shock, which are assumed same for all shocks and shock effects are considered independent., we have the probability of survival of k shocks,

$$\overline{H(t)} = e^{-\lambda t(1-q)}$$

The average life of such a device

$$E(T) = \int_0^\infty (\overline{H(t)})dt$$
$$= \int_0^\infty e^{-\lambda t(1-q)}dt$$
$$=[\lambda(1-q)]^{-1} \tag{4.2.9}$$

Using this information we can formulate the loss due to replacement of such a device at time 't'. Here we wish to obtain the optimal replacement age minimizing the loss associated with the replacement. In order to reduce the risk of failure we assume that the system be replaced at or before its expected life. If the device is replaced at time 't', the remaining operational time is $E(T) - t$. If the device is replaced at the starting stage of its life, it may incure a severe loss due to the unused operational time of the system. So to reflect the big loss in the early replacement, it is natural to model the loss in this case while replacing the device at time 't', in terms of the square of the remaining operational time, $(E(T)-t)^2$. Hence the loss due to the replacement at time 't' can be taken as $C_1(E(T)-t)^2$,where C_1 is the constant of proportionality reflecting the degree of severity due to the early replacement. As we choose 't' near $E(T)$,the loss of the above type will be minimal, but here the risk of failure of the device before replacement is greater. Moreover as time goes on the level of the functioning of the device (by many shocks acted on it), may be decreased. In such situations the loss may be treated as directly proportional to the cumulative hazard of the device at time 't'. So such a loss at 't' can be represented as $C_2(-log_e S(t))$, where C_2 represents the constant of proportionality which reflects the degree of hazard and $S(t)$ is the survival function of the device.

Combining these two types of losses at time 't', we can formulate a natural loss function due to the replacement of the device at time t, as

$$L(t) = C_1(E(T) - t)^2 + C_2(-log_e S(t)) \qquad (4.2.10)$$

One may think of alternative loss functions associated with the present problem in a meaningful manner and seek the optimal replacement age accordingly.

In the present context, the time point 't' which minimizes $L(t)$ is the optimal replacement age of the device, corresponds to the given C_1 and C_2.

Obviously the nature of the loss function heavily depends on the life distribution, particularly on the form of its survival functions. Also under the Poisson shocks, the nature of the shock probabilities decides the nature of the life distribution of the device. When shock probabilities are DPH, then we get closed form expression for the loss functions, which allow deriving optimum replace age of the device. We discuss these cases separately in the sequel.

Here in the case of device with geometric shock probabilities,

$$L(t) = C_1[\frac{1}{\lambda(1-q)} - t]^2 + C_2[-log_e e^{-\lambda t(1-q)}] \qquad (4.2.11)$$

$$L(t) = C_1[\lambda(1-q))^{-1} - t]^2 + C_2[\lambda t(1-q)] \qquad (4.2.12)$$

The 't' value which minimises the quadratic function $L(t)$ can be derived as,

$$t^* = \frac{1}{\lambda(1-q)}[1 - \frac{C_2\lambda^2(1-q)^2}{2C_1}] \quad (4.2.13)$$

As an illustration we consider a case with $\lambda = 2$,$q = 0.4$, $C_1 = 10$ and $C_2 = 8$. From the graph of $L(t)$ (Fig.No.1) ,plotted using the given parameter values, the optimal replacement age is the 0.36 units of time.

Poisson shocks with Discrete Phase type shock probabilities

There are several real life situations where one never knows the precise form of distribution for the shock probabilities $\{p_k\}$. So in such general situations we can make use of the phase type modelling to solve the problem. The class of phase type distribution is dense, and hence any distribution with positive support can be approximated arbitrarily close by a phase type distribution. See Neuts(1981). In such cases Manoharan et. al. (1992) established that the survival probability at time t of the device can be represented as

$$\overline{H(t)} = \alpha e^{\lambda(T-I)t}\underline{e}, \quad (4.2.14)$$

where λ is the rate of Poisson shocks. So the life distribution of the device can be represented as a Continuous Phase Type (CPH) Distribution with parameters α and $\lambda(T-I)$

Therefore the expected life of the device

$$E(T) = -\alpha[\lambda(T-I)]^{-1}\underline{e} \quad (4.2.15)$$

Hence the Loss function $L(t)$ has the form;

$$L(t) = C_1(-\alpha[\lambda(T-I)]^{-1}e - t)^2 + C_2(-log_e(\alpha e^{\lambda(T-I)t}e)) \qquad (4.2.16)$$

By standard calculus method, the optimal replacement age is obtained as,

$$t^* = \frac{C_2\lambda(T-I)}{2C_1} - \alpha[\lambda(T-I)]^{-1}\underline{e} \qquad (4.2.17)$$

To illustrate the above method, first we consider the same problem discussed in the previous section. We may represent the geometric shock probability $\{p_k\}$ as a discrete phase type variate with parameters $\alpha = (1)$ and $T = (q)$. The lifetime of the device $T \sim CPH(\alpha, \lambda(T-I))$ so here $T \sim CPH((1), \lambda(q-1))$

For the given set of parameter values $\lambda = 2, q = 0.4, C_1 = 10, C_2 = 8$, the optimal replacement age as given by (4.2.13) is calculated as $t^* = 0.36$ time units which again agrees with the former illustrated example.

Next we consider a non-trivial example with DPH shock probabilities. The device is suffering from homogenous Poisson shocks with parameter λ=2 and, the shock probabilities following DPH with parameters $\alpha = (1,0)$ and

$$T = \begin{bmatrix} 0.65 & 0.1 \\ 0.15 & 0.8 \end{bmatrix}.$$

Hence the lifetime of the device $T \sim CPH(\alpha, \lambda(T-I))$. For $C_1 = 1$ and $C_2 = 7$, the optimal replacement age of the device is computed to be 1.3 time units. Figure 2 shows the loss function and the optimal replacement of the device.

Remark 4.2.1 *In the case of an arbitrary distribution with the positive support for the shock probabilities p_k, one may approximate it as closely as possible by a DPH distribution. Hence the foregoing methods can be followed to obtain the optimal replacement age in such a general situation.*

The gain due to age replacement

Here while finding the gain due to replacement we are taking the cost of preventive replacement as $V + L(t)$, where V is the price of the new device and $L(t)$ is the cost due to the replacement at time $'t'$ and the cost of failure replacement as M. Hence the long run cost per unit time in this case of age replacement is,

$$C_a = \frac{[(V + L(t)](1 - F(t)) + [M](F(t))}{\int_0^t (1 - F(t))dt} \qquad (4.2.18)$$

and the expected long run cost per unit time in the case of failure replacement is,

$$C_f = \frac{M \times \text{Expected number of renewals in (o,t)}}{t} \qquad (4.2.19)$$

This can be calculated as described earlier.

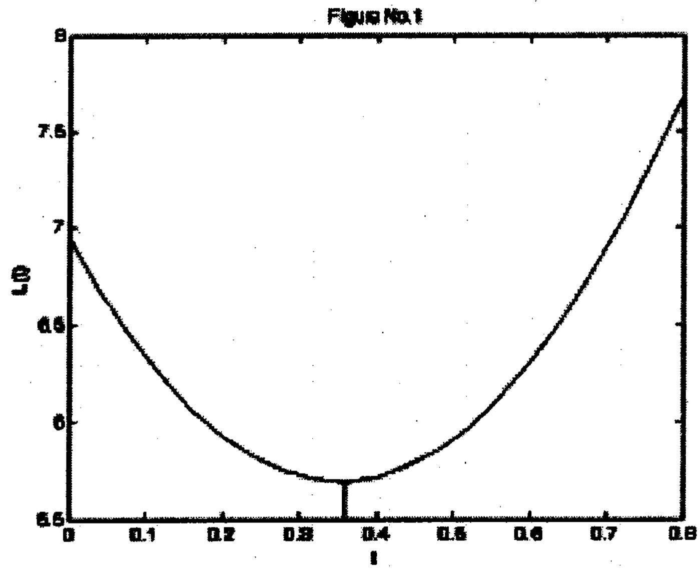
To confirm the supremacy of the method, we re-consider the case of a device with homiogenous shocks and geometric shock probabilities described earlier. We got the optimum replacement age as 0.36 time units and the loss due to this replacement $L(t)$ as 5.6964 and suppose for instance, V= 50 and M = 5000. Then we are getting $C_a = 6.1237 \times 10^3$ and $C_f = 1.959 \times 10^4$. Therefore $G = 1.3467 \times 10^4$.

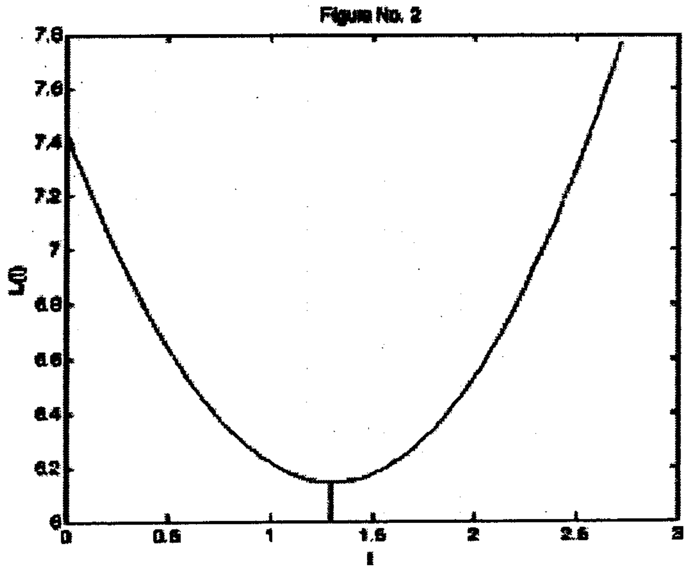
In another case of a device with life distribution $CPH(\alpha, T)$ where $\alpha = (0.6, 0.4)$ and

$$T = \begin{bmatrix} -0.6 & 0.2 \\ 0.4 & -0.4 \end{bmatrix}$$

with the optimal replacement age 4 time units, for v=50 and M=3500 , we get $C_a = 823.0761$ and $C_f = 1.0383 \times 10^3$, therefore $G = 9.5596 \times 10^3$.

4.3 Figures:





MULTISTATE DETERIORATING SYSTEM SUBJECT TO POISSON SHOCKS

Aneesh Kumar K. “Phase type modelling in reliability and survival analysis” Thesis. Department of Ststistics , University of Calicut, 2007

Chapter 5

MULTISTATE DETERIORATING SYSTEM SUBJECT TO POISSON SHOCKS

[1]Contents of this chapter is accepted for publication in,

The Proceedings of the National Seminar on Stochastic Modelling,Distribution Theory and Order Statistics, March 2006, University of Kerala.

5.1 Introduction

A fundamental concept in reliability analysis is that of a binary system, in which the system or component are considered to be in one of the two states-functioning or failed. But it is natural to think that the dichotomized modelling of the system over simplifies the reality. In many real life situations, however, the systems are better modelled by allowing a range of levels of performance, from perfect functioning down to complete failure. In these situations some kind of multistate model is essential. Very early the topic was discussed by Barlow and Wu(1978), El-Newihi et al (1978) etc. And a discussion on multistate reliability is performed by Grifith(1980). The multistate reliability theory can handle situation in which the system and its components have a range of performance levels from perfect functioning to complete failure. Manoharan and Joseph(2004) considered a multistate system with different performance level and suggested an optimal replacement schedule for the system. Because of performance degradation is very common in industrial products, it is important to develop the reliability theory for multistate systems. Here we are considering a multistate system subject to Poisson shocks.

In real life situations we can see that many systems reaches the end of its life span by moving through different states of their performance level, from best to worse. There may be some factors which affect the system adversely and lead the system to failure. Depending on the environment in which the system works, the adverse effects on the system (which we can call as shocks) may vary and thereby the lifespan of the system may become short. In such situations the question is about the life of the system. Narasimhalu and Prasad (1997), studied such a system with accumulated damage. Here we model the life of the system in such an environment with the help of phase type distribution.

Consider the problem of finding the lifetime of a multistate system which is subjected to Poisson shocks at its working environment. The distribution of the magnitude of shocks that would act on the system is known and the system goes from one working state to the next deteriorated state and finally to the failure state due to the shocks. In such cases the proposed method can be used to find the life distribution of the system. We consider the shock probabilities as having discrete phase type (DPH) distribution and obtain the life distribution of the system as a continuous phase type distribution. Hence the reliability, hazard rate etc., of the system can be discussed.

The results following have applications in several medical situations, for instance, in the study of survival experience of patients with cerebral hemorrhage.

5.2 The model

Consider a multistate system. The system consists of n different working states. A new system which is functioning perfectly is said to be in state 1. State 2 is the next deteriorated functioning state of the system and the n^{th} state is the complete failure state of the system. The system is subjected to homogenous Poisson shocks with parameter λ. The system is getting deteriorated and reaches its failure state due to the shocks acted on the system. The magnitude of the shock acted on the system determines the severity of the deterioration towards the failure of the system. Any shock can deteriorate the condition of the performance level of the system. If there are no shocks acted on the system, it is assumed to continue in its current working state. Let Z denote the magnitude of the shock acted on the system. When one designs a system, he knows the capacity of the system to suffer shocks. The quantum of shocks that will shift the performance level of the system is as follows:

The quantum of shocks less than or equal to $'k'$ shifts the performance level of the system by one unit. Specifically, for the system working in i^{th} state where $(1 \leq i \leq n-1)$, if the magnitude of the shock Z, acted on the system is such that, $0 < Z \leq k$, system get deteriorated to $(i+1)^{th}$ state. If $k < Z \leq 2k$, it get deteriorated to $(i+2)^{th}$ state.

In general, if $(r-1)k < Z \leq rk$, for $r = 1, 2, \ldots, (n-i-1)$ system gets deteriorated to $(i+r)^{th}$ state and if $Z \geq (n-i-1)k$, the system directly jumps to the failure state from the i^{th} state. There is no maintenance facility available to improve the current working state of the system. So if the system in in the i^{th} state, according to the shock magnitude Z, it only moves to a state j, where $i < j \leq n$.

Let p_i denote $Pr[(i-1)k < Z \leq ik]$ for $i = 1, 2, \ldots, (n-2)$. and $\beta_i = Pr[Z > (n-i-1)k]$, for $i = 1, 2, \ldots, n-2$.

Note that,

$$\beta_{n-1} = Pr[Z > 0] = 1$$

Define Y_n = The working state of the system after the n^{th} shock acted on it, where $n = 1, 2, \ldots$. Then $\{Y_n\}$ is a Markov chain with initial probability vector $\underline{\alpha} = (1, 0, 0, \ldots, 0)$ and the transition probability matrix

$$P = \begin{bmatrix} 0 & p_1 & p_2 & p_3 & \cdots & p_{n-2} & \beta_1 \\ 0 & 0 & p_1 & p_2 & \cdots & p_{n-3} & \beta_2 \\ 0 & 0 & 0 & p_1 & \cdots & p_{n-4} & \beta_3 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & p_1 & \beta_{n-2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

we may write the block partitioned form as,

$$P = \begin{bmatrix} T & \underline{\beta} \\ \underline{0} & 1 \end{bmatrix},$$

where,

$$T = \begin{bmatrix} 0 & p_1 & p_2 & p_3 & \dots & p_{n-2} \\ 0 & 0 & p_1 & p_2 & \dots & p_{n-3} \\ 0 & 0 & 0 & p_1 & \dots & p_{n-4} \\ . & . & . & \dots & . & . \\ . & . & . & \dots & . & . \\ 0 & 0 & 0 & 0 & \dots & p_1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

$\underline{0} = (0,0,0,\dots 0)$, and, $\underline{\beta} = \underline{e} - T\underline{e}$, given $\underline{e} = (1,1,1,\dots 1)_{(n-1)\times 1}$.

Let X denote the number of transitions required for this Markov chain to get into the absorbing state n. It is to be noted that the transition in the working phase of the system is solely due to the shocks acting on it and each shock will result in a state change however small its magnitude be. Therefore the number of transitions required for the Markov chain Y_n to get in to the absorbing state is equal to the total number of shocks acted on the system which leads the system to failure. This implies P(X=k) = P(Number of shocks acted on the system leading the system to failure=k) Notice that, the probability distribution of the number of transitions required for an absorbing Markov chain to get into its failure state is a discrete phase type (DPH) distribution with parameters $\underline{\alpha}$ and T.

Obviously we get, the number of shocks to be acted on the system to its failure follows the same DPH. That is the shock probabilities a_k, the probability of the system fails by k^{th}, follows $DPH(\underline{\alpha}, T)$.

In our model the arrival of the shocks is considered as a counting process, where, $N(t)$, the number of shocks in the time interval (0,t) is taken as a Poisson process with parameter λ.

Let $\overline{A_k}$ be the probability of surviving k shocks by the system, for k = 0,1,2, ... where $1 = \overline{A_0} \geq \overline{A_1} \geq \overline{A_2} \ldots$

Here ,

$$\overline{A_k} = \sum_{n=k+1}^{\infty} \alpha T^{n-1}\beta \quad (\text{Since } X \sim DPH(\underline{\alpha}, T)) \tag{5.2.1}$$

$$= \underline{\alpha} T^k \underline{e}_{(n-1)} \quad (\text{Since } T\underline{e}_{(n-1)} + \beta = \underline{e}_{(n-1)}) \tag{5.2.2}$$

Now from Manoharan et al (1992) , the survival function $H(t)$ of the system can be obtained as

$$H(t) = \sum_{k=0}^{\infty} P(N(t) = k)\overline{A_k} \tag{5.2.3}$$

$$= \underline{\alpha} e^{\lambda(T-I)t} \underline{e}_{(n-1)} \tag{5.2.4}$$

which is the survival function of a random variable following continuous phase type distribution with parameters $(\underline{\alpha}, \lambda(T-I))$,where I is the identity matrix. So the lifetime of the system (let it be denoted by Y) follow CPH$(\underline{\alpha}, \lambda(T-I))$ That is $Y \sim CPH(\underline{\alpha}, \lambda(T-I))$. This implies that the arrival rate of Poisson shocks and distribution of the shock magnitude 'Z' completely specifies the lifetime of the multistate system considered above.

5.3 System under Poisson shocks with exponential shock magnitude

Consider a system suffering from Poisson shocks with arrival rate λ_1 . $N(t)$ denote the number of shock arrivals in (0,t), then.,

$P(N(t) = n) = \frac{e^{-\lambda_1 t}(\lambda_1 t)^n}{n!}$; n = 0,1,2, ...

Let the magnitude of shocks 'Z' is distributed as exponential random variable with probability density function $f(z) = \lambda_2 e^{-\lambda_2 z}, z > 0$. Here we are defining n functioning states to the system. According to the magnitude of the shock acted on the system, it will change the states of the system from the perfect functioning to failure as described earlier. The value of 'k' is selected for the system according to the capacity of the system to work under shocks. By using $f(z) = \lambda_2 e^{-\lambda_2 z}, z > 0$, we can find the values of p_1,p_2, ...for the selected value of 'k' and can complete the matrix P. Hence we can identify the matrix T from P. Since the shock arrival follows Poisson distribution we can set the life distribution of the system as a CPH random variable with parameters$(\underline{\alpha}, \lambda_1(T - I))$, that is the life of the system $Y \sim CPH(\underline{\alpha}, \lambda_1(T - I))$.

5.4 Mathematical Illustration

For a mathematical illustration of the above description and for finding the reliability characteristics, we assume a system having six different phases including one absorbing phase. Let the magnitude of shocks acting on the system is distributed as $f(z) = 2e^{-2z}, z > 0$. Also let the shock arrival be Poisson with parameter $\lambda_1 = 3$. Under this environment, for a system with value of k = 0.25 as the tolerance period of the magnitude of shocks for each phase, we can get the matrix P as

$$P = \begin{bmatrix} 0 & 0.3935 & 0.2387 & 0.1447 & 0.0878 & 0.0878 \\ 0 & 0 & 0.3975 & 0.2387 & 0.1447 & 0.2231 \\ 0 & 0 & 0 & 0.3975 & 0.2387 & 0.3679 \\ 0 & 0 & 0 & 0 & 0.3975 & 0.6065 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

So the life Y of the system is distributed as $CPH(\underline{\alpha}, 3(T-I))$;

where $\underline{\alpha} = (1, 0, 0, 0, 0)$ and,

$$T = \begin{bmatrix} 0 & 0.3935 & 0.2387 & 0.1447 & 0.0878 \\ 0 & 0 & 0.3975 & 0.2387 & 0.1447 \\ 0 & 0 & 0 & 0.3975 & 0.2387 \\ 0 & 0 & 0 & 0 & 0.3975 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence the reliability of the system at the time point t,

$$R(t) = \underline{\alpha} e^{3(T-I)t} \underline{e}$$

and the hazard function,

$$h(t) = \frac{\underline{\alpha} e^{3(T-I)t} \underline{\beta}}{\underline{\alpha} e^{3(T-I)t} \underline{e}}.$$

Using MATLAB, we obtain the plot of the reliability function of the system as shown in Figure No. 1.

Remark 5.4.1 *If the shock is a mixture of Poisson process, due to the environmental factors, then also the system life can be modelled as following a phase type distribution. When $N(t)$ is a mixture of Poisson process with mixing density $G(\lambda)$ having finite support $(\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_k)$ with mixing density $(q_1, q_2, q_3, \ldots q_k)$, then the life of the system follows $CPH(\alpha^*, T^*)$; where $\alpha^* = (q_1\alpha, q_2\alpha, q_3\alpha, \ldots q_k\alpha)$ and $T^* = diag(\lambda_1(T - I), \ldots \lambda_k(T - I))$.*

5.5 System under Poisson shocks with unknown distribution for shock magnitudes

The above cited method for finding the life and reliability of a multistate system subjected to Poisson shocks acted can be easily applied to practical situations, where the distribution of the shock magnitude is unknown also. In such situations the empirical estimates of $p_i^{\prime}s$ can be had from the data on the shock magnitude that would act on the system as follows,

$$\hat{p}_i = \frac{\text{Number of shocks with magnitude in the interval (k(i - 1) ,ik]}}{n} \qquad (5.5.1)$$

Suppose we are interested in the reliability of a system which is installed at a certain place. The environmental conditions of different places are different. We have knowledge about the system and its tolerance level towards the shock magnitude. Consider the possible shocks, that may act on the system at the proposed place. Suppose we have the data on the magnitude of a series of large number n of shocks which would act on the system at these places. Using this data, we can get empirical estimates of elements $p_1, p_2, \ldots$ of the transition probability matrix corresponding to the Markov chain associated with the different phases of the system as in (5.5.1)

To illustrate the method, we shall generate data having 1000 values on shock magnitudes of Poisson shocks with rate $\lambda = 2$ that would act on a system. We estimate the reliability and hazard rate of the system as follows. The system is considered having 6 states where the state 6 is the failure state and we assume the value of k = 0.2. The classified data on shock magnitude is shown below.

$Shock Magnitude$	$No. of shocks$
$0 < Z \leq 0.2$	634
$0.2 < Z \leq 0.4$	238
$0.4 < Z \leq 0.6$	89
$0.6 < Z \leq 0.8$	28
$Z > 0.8$	11
Total	1000

The transition probabilities are empirically estimated as $\hat{p}_1 = 0.634$, $\hat{p}_2 = 0.238$,$\hat{p}_3 = 0.089$, $\hat{p}_4 = 0.028$, and $\hat{p}_5 = 0.11$. Hence the life of the system follows $CPH[\underline{\alpha}, \lambda(T - I)]$

where, $\underline{\alpha} = (1, 0, 0, 0, 0)$ and the matrix T is,

$$T = \begin{bmatrix} 0 & 0.634 & 0.238 & 0.089 & 0.028 \\ 0 & 0 & 0.634 & 0.238 & 0.089 \\ 0 & 0 & 0 & 0.634 & 0.238 \\ 0 & 0 & 0 & 0 & 0.634 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The reliability and hazard plots of this system are plotted in Figure No.2.

Remark 5.5.1 *The problems of interest which employs reliability and hazard rate of the system, can be solved hopefully by reaching the life distribution and the plots of reliability and hazard function of the system.*

5.6 Figures:

Figure No. 1

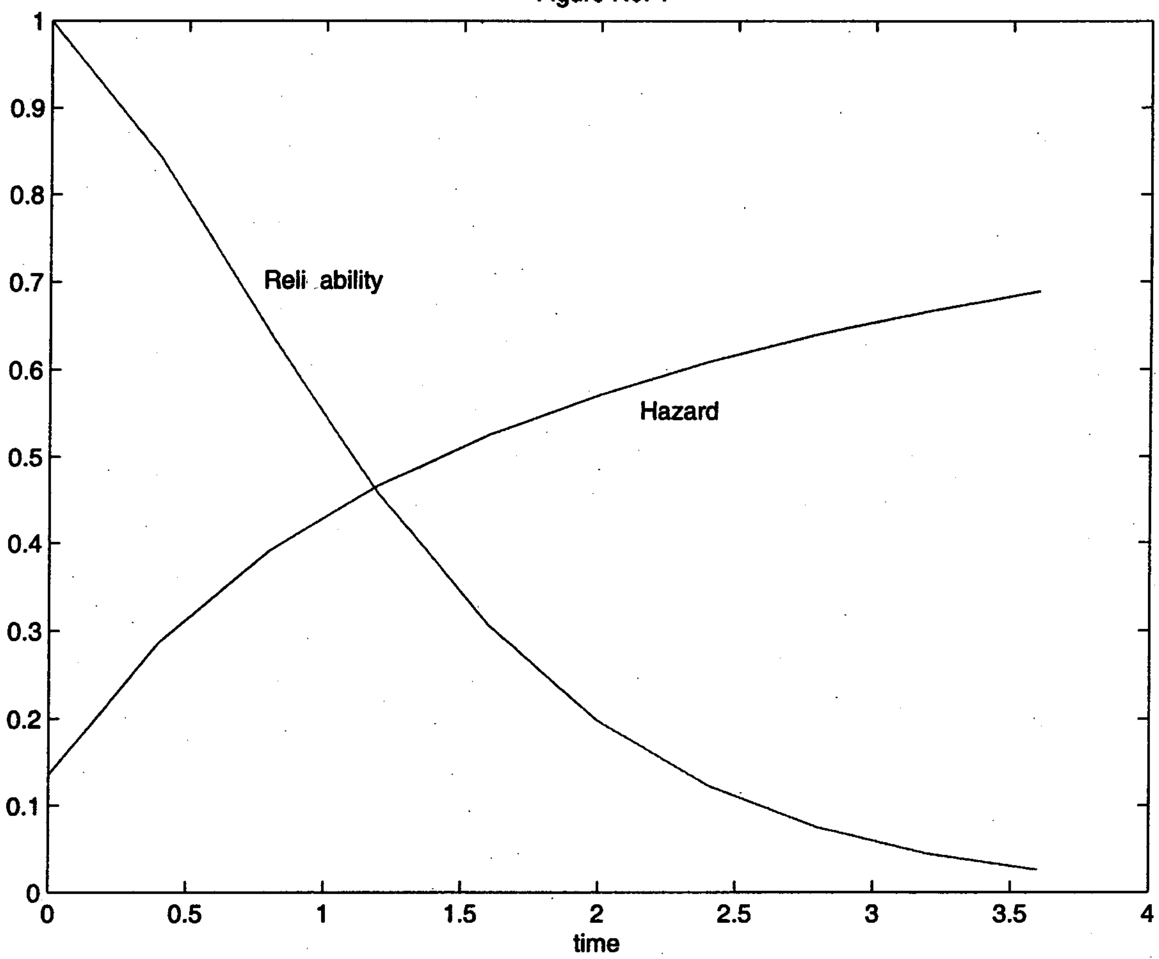
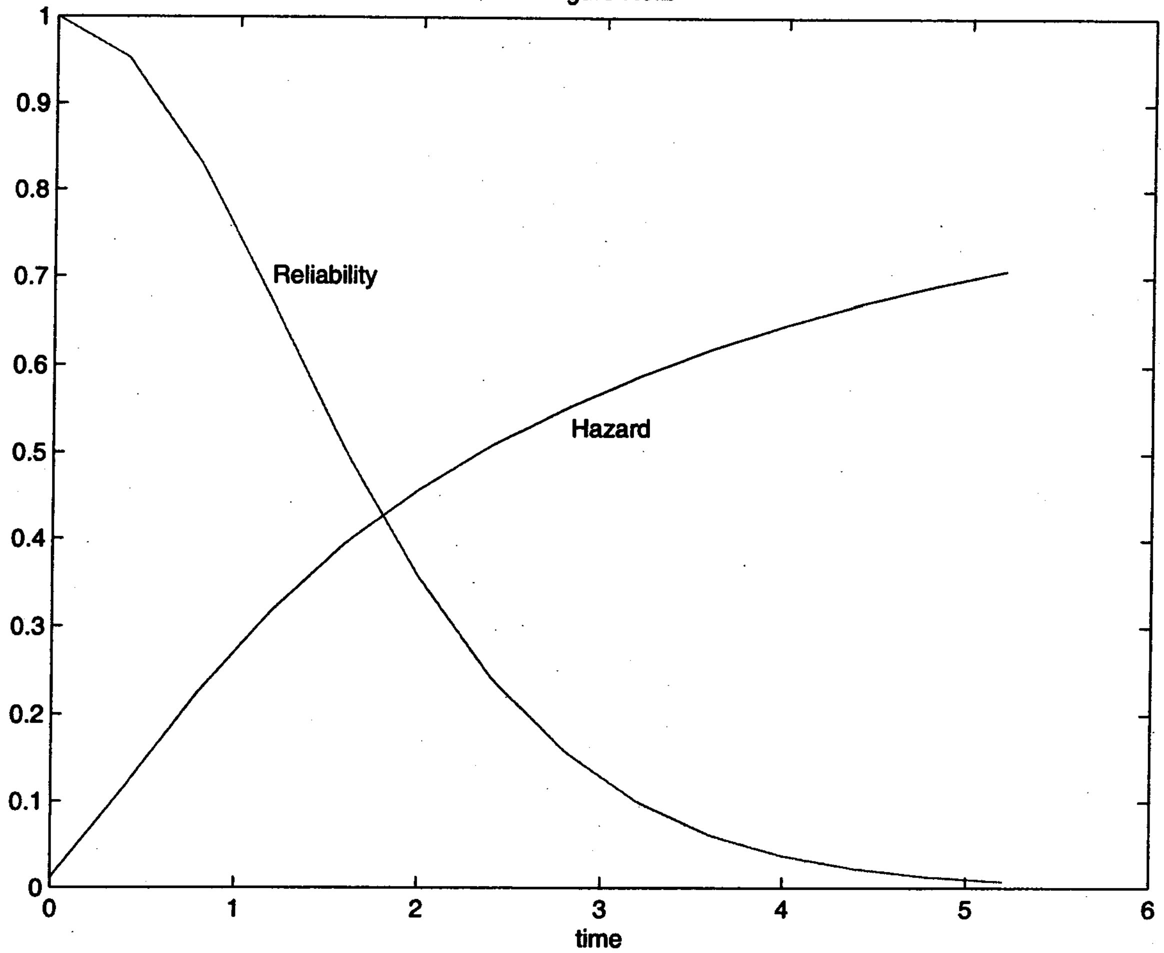


Figure No.2



MODELLING SEED GERMINATION - AN APPLICATION OF PHASE TYPE DISTRIBUTION

Aneesh Kumar K. “Phase type modelling in reliability and survival analysis”
Thesis. Department of Ststistics , University of Calicut, 2007

Chapter 6

MODELLING SEED GERMINATION - AN APPLICATION OF PHASE TYPE DISTRIBUTION

[1] Contents of this chapter is accepted for publication in,

Advances and Applications in Statistics, (2007)

6.1 Introduction

Stochastic modelling of seed germination is a rather unexplored area of research and finds great application and use in plant sciences. The information on specific stages of germination process may help one to have some precautionary or remedial measures to check them. Statistical Inference on the underlying process based on censored or incomplete observations remains to be an ever challenging one in such investigations.

Seed germination is the resumption of active growth of the embryo that results in the rupture of the seed coat and the emergence of the young plant. Seed germination is a complex process involving many phases. Some seeds are capable of germination soon after fertilization, while other may be dormant and require an extended rest period for germination. The major events occurring in seed germination are water imbibition, enzyme activation, initiation of embryo growth, rupture of seed coat and emergence of the seedling and finally establishment of the seedling. For the fundamental notions of seed germination, one may refer to Copeland(1988). Shafili and Price(2001) discuss on modelling the time of seed germination at various level of temperature and on the estimation of cardinal temperature. Michael et.al(2004) reviews methods for fitting a range of models to censored seed germination data and recommends adoption of a probability -based model for the time to seed germination. The inverse normal distribution, is considered as a distribution for the time to germination.

In this chapter we model seed germination time under phase type distribution assumption. We are considering the time required to get an established seedling for plantation, possibly for any variety of seed under standard or controlled conditions. The different stages in the growth of the plant is known. In this study we are considering seeds coming under epigeal germination process. Under epigeal germination the cotyledon are raised above from soil and they continue to provide nutrition support to the growing plants. In its growth, the hypocotyl begins to elongate in an arch which breaks through the soil, pulling the cotyledon and enclosed plumule through the ground and projecting them into the air. Afterwards the cotyledon open and plumule growth continues and the exhausted cotyledon wither and fall to ground, then the leaves develop and seedling get ready to establish. It may help one interested in the physiology of seeds to gather more information on the growth stages like the rate of growth, the hazard nature of the seeds in its different growth stages.

For the description of the model we are considering four noticeable stages of the growth of the plant. The seed is considered in the 0^{th} stage. When the hypocotyl coming out of the soil, it is said to be in the first stage of the growth. When the hypocotyl is partially developed, that is about 2 cms. the plant is considered in the second stage of its growth. The opening of the cotyledon and thereby the appearance of the leaves is the third and and the seedling with fully expanded leaves considered mature for plantation and this stage is taken as the fourth stage of the growth. It is observed that the growth rate of the seedling in different stages are different and so are the distribution of the time to cover the stages also. For the probabilistic modelling of the occupation time in these stages, a more flexible class of probability distribution on the positive real line is desirable, which allows better inferential capability. Here we use a highly versatile class of probability distribution, that of Phase type distibution - introduced by Neuts(1975).

The convolution property (2.5.4) of the Phase type distribution that is used in the model description. In the parameter estimation of phase type distribution which involving censored observations, we use the EMpht-programme described in the section (2.8) .

6.2 The Model for Seed Germination

Since any distribution on positive support can be approximated by phase type distribution, we are making use of the method of phase type distribution for the probability distribution of time to reach the seedling to a stage from the previous. In between each stage transition existence of many transient phases can be assumed. When we are considering the time to reach from one stage to the next nearest stage, the starting stage is the initial phase and the next stage is the absorbing phase. The time required to reach from the initial phase to the absorbing phase through many phases in between can be modeled as a continuous phase type distribution (CPH).

This modelling can be done for each stage transition. Let X_1 denote the time to reach first stage, X_2, X_3 and $X_4$are the same for second, third and the fourth stages respectively. Though the quality of seeds effect the growth rate of seed in each stage, there are many random environmental factors which may affect the growth rate of the seed in all stages independently. So one assume that $X_i's$ are independent. Under CPH assumption, let X_i follow $CPH(\alpha_i, T_i)_m$ for $i = 1, 2, 3, 4$. Then due to the convolution theorem for CPH distribution, it follows that the total germination period, $X = X_1 + X_2 + X_3 + X_4$ follow $CPH(\gamma, L)$, where $\gamma = (\alpha_1, \alpha_{1,m+1}\alpha_2, \alpha_{1,m+1}\alpha_{2,m+1}\alpha_3, \alpha_{1,m+1}\alpha_{2,m+1}\alpha_{3,m+1}\alpha_4)$ and

$$L = \begin{bmatrix} L_2 & {L_2}^0\alpha_4 \\ \underline{0} & T_4 \end{bmatrix}$$

where,

$$L_2 = \begin{bmatrix} L_1 & {L_1}^0\alpha_3 \\ \underline{0} & T_3 \end{bmatrix}$$

and

$$L_1 = \begin{bmatrix} T_1 & \underline{T_1}^0\alpha_2 \\ \underline{0} & T_2 \end{bmatrix}$$

Here for the i^{th} stage, the probability of more than 't' hours is taken to reach the next stage is given by the survival function $P(X_i > t)$, and since X_i follow

$CPH(\alpha_i, T_i)$,

$$P(X_i > t) = \alpha_i e^{T_i t} \underline{e}.$$

where, $\underline{e}$ is a unit vector of dimension m x 1.

$P(X_i > t)$ measures the latent reliability of the i^{th} stage and is a measure of the growth of the plant in that stage.

Probability of instantaneous reach to the next nearest stage from a stage can be denoted by $h(t)$ and is defined by

$$h(t) = \frac{f(x)}{1 - F(x)}.$$

For the random variable X_i , $h_{X_i}(t)$ is the instantaneous growth rate of the plant in the time duration of its growth from $(i-1)^{th}$ stage to i^{th} stage.

Also $E(X_i)$ gives us about the expected waiting time to reach stage i from $(i-1)^{th}$ stage, while $E(X)$ gives the expected total time for getting a mature plant for ready to replant from its seed.

6.3 Estimation of Distributions

The parameters (α_i, T_i) and, thereby the distribution of X_i can be estimated using EMpht algorithm for fitting phase type distribution. Using the observed times to reach the absorbing stage, one can make use of the algorithm for fitting corresponding phase type distribution. The procedure can be used for data set including censored observations also. In the present context, we are observing the seeds and noting their time to reach their first stage of growth. It is not sure that all the seeds will survive to reach this stage within a finite time period and therefore it is worthless trying to collect the observations from all the seeds. In such cases we are satisfying with the data from a reasonable higher percentage of seeds and the data for other seeds are considered as censored observations at the time point where we collected the last observation. Since we are interested in the time taken by the seed to reach each stage from the previous, we are not considering the decayed seeds in between. Using these data the parameters (α_1, T_1) for the phase type representation of X_1 is estimated. For fitting phase type distribution to X_2, we are following the seeds reached the first stage and observing their time to reach the second stage, but not for all but for a major percentage and others are considered as censored due to the early explained reason. Similarly the distribution of all $X_i's$ are fitted using the data observed.

The data on the time of seed decay, if any, can be used to model the distribution of seed life. The survival chance of the seeds to a mature plant which is ready to replant can be found, using the decay time of the decayed seeds. We are noting the decay time of all the seeds decayed up to our censored time point for the last stage. Out of the total of N seeds, let n be decayed and for the others the decay times are right censored at various time points. On observing the seeds in seed stage to its first stage of growth the seeds are under soil and we are observing for the appearance of hypocotyl for a major percentage of the seed planted and the others are considered as censored. Many of the seeds may decayed under soil, but unseen. Then later on a close observation, it is found that the delay in growth of almost 40 percentage of the seeds which are considered censored is due to its decay under soil within that censoring time. So when considering the decay time of that 40 percentage of seeds, we assume that their decay is happened at the mid of the time duration from seed stage to the censoring time for the stage. The resulting observations are then summarised. Using the EMpht algorithm the distribution of the life time of the seeds can be found in phase type distribution form. From here we can plot the graph of the survival function of the seeds and the hazard plot also. The plots can be used for finding the hazardous nature of the seed. Modelling of the life distribution of the seed in phase type form in each stage is done using the time of decay of the seeds which occurred at that stage and the censoring times. A relative study on the hazardous nature of the seed in each of its growth stage can also be obtained by ploting the hazard curve of the life distribution in each stage.

6.4 Experimental Validation

For an illustration of the above method, we are making use of an experimental data (see Table - I)obtained on 100 numbers of the seed of green peas (*Pisum sativum*) observed for germination and to become a mature plant for plantation under standard conditions. The seeds are continuously observed and time required to reach

each stage of its growth- which are already mentioned in the description- from its previous growth stage is noted in hour and rounded to the nearest integer. All seeds are planted for germination at the same time and the time for the appearance of hypocotyl is collected for about 80 percentage of the seeds and the others are considered censored at the highest time point observed. The phase type form of the distribution for the time required to reach this first stage from the seed stage is estimated using the data collected and using the EMpht algorithm. A general Coxian form of the phase type distribution is fitted here,because when we are considering the growth a backward movement in phases is not reasonable. It is observed that this time duration denoted by X_1 follows phase type distibution with parameters $\alpha_1 = (1, 0)$ and

$$T_1 = \begin{bmatrix} -0.057735 & 0.057735 \\ 0 & -0.057735 \end{bmatrix}$$

In a similar way the distributions of the time duration required for each stages,X_2,

X_3 and X_4 are estimated as in the general Coxian for of phase type distribution using the observed data and the parameters estimated are as follows: for X_2 , $\alpha_2 = (1, 0)$

$$T_2 = \begin{bmatrix} -0.053238 & 0.053238 \\ 0 & -0.053258 \end{bmatrix}$$

for X_3, $\alpha_3 = (1,0)$

$$T_3 = \begin{bmatrix} -0.06366 & 0.06366 \\ 0 & -0.06366 \end{bmatrix}$$

And for X_4 $\alpha_4 = (1,0)$

$$T_4 = \begin{bmatrix} -0.034885 & 0.034885 \\ 0 & 0.034885 \end{bmatrix}$$

Then $X = X_1 + X_2 + X_3 + X_4$ follows phase type distribution with parameters

$\alpha = (1,0,0,0,0,0,0,0)$ and

$$T = \begin{bmatrix} -0.0577 & 0.0577 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.0577 & 0.0577 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.0532 & 00532 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.0532 & 0.0532 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.0637 & 0.0637 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0-0.0637 & 0.0637 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.0349 & 0.0349 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0349 \end{bmatrix}$$

The expected time durations $E(X_1)$,$E(X_2)$,$E(X_3)$,$E(X_4)$ are calculated using

$$E(T_i) = -\alpha_i T^{-1} \underline{e}.$$

The values are calculated as 34.64, 37.56,31.41 and 58.52 hours respectively. And $E(X) = 160.9563$. *Figure*1 gives the nature of instantaneous growth rate for each stages of growth and it shows that the growth rate is highest is from stage

2 to stage 3 and is least for the stage 3 to stage 4. *Figure*2 shows the curves regarding the probability of the chance at next stage at a time t starting from a particular stage.

Using the observed decay times, the life distribution of the seed (see Table-II) is derived and the life distribution of the seed in each stage is also estimated. The life distributions are estimated in a general phase type form and the life of the seed up to mature stage L, is distributed as phase type with parameters

$$\beta = (0.312411, 0.687589, 0)$$

$$S = \begin{bmatrix} -0.032552 & 0.004845 & 0.027707 \\ 0.044172 & -0.065707 & 0.021535 \\ 0.006152 & 0.012213 & -0.030373 \end{bmatrix}$$

The life from seed stage to first stage, L_1; from first stage to second, L_2; from second stage to third, L_3; and from the third stage to the last, L_4 are respectively distributed as general phase type with parameters (β_1, S_1), (β_2, S_2), (β_3, S_3), (β_4, S_4) They are estimated as; $\beta_1 = (0.999965, 0.000035, 0)$

$$S_1 = \begin{bmatrix} -0.167049 & 0.167044 & 0.000005 \\ 0.017308 & -0.131828 & 0.11452 \\ 0.027979 & 0.058696 & -0.11937 \end{bmatrix}$$

$\beta_2 = (0.001334, 0.800287, 0.198378)$

$$S_2 = \begin{bmatrix} -0.084882 & 0.017825 & 0.05975 \\ 0.046926 & -0.166635 & 0.119709 \\ 0.151992 & 0.051021 & -0.203013 \end{bmatrix}$$

$\beta_3 = (0.804312, 0.195688, 0.000001)$

$$S_3 = \begin{bmatrix} -0.358696 & 0.212524 & 0.146172 \\ 0.240711 & -0.507013 & 0.266303 \\ 0.128051 & 0.150568 & -0.311636 \end{bmatrix}$$

and $\beta_4 = (0.203247, 0.394742, 0.402011)$

$$S_4 = \begin{bmatrix} -0.253126 & 0.166938 & 0.083404 \\ 0.105748 & -0.205621 & 0.096256 \\ 0.115563 & 0.161631 & -0.292831 \end{bmatrix}$$

Figure 3 exhibits the nature of the hazard of the seed from its seed stage to its final stage. Figure 4 shows how the hazard function differ for the different growth stage of the seed and it reveals that the hazard is highest for the time from seed stage to the first stage and it is least for the growth from the first stage to second.

• Discussion:

In this paper our interest centers on the process of seed germination which involves many intermediate stages. The time required for germination as well as the sojourn times in the growth stages have been studied under a general setting. Though the number of stages is taken as four in the present study, the methodology can be carried over to any number of stages in the germination process. Also quite important information on the growth rate and hazard rate at each intermediate stages of the germination process are obtained and illustrated through an experimental data. The choice of phase type distribution in the model is a natural one as it does not loose any generality. The properties of this versatile class of distributions is exploited in further gathering information on the seed germination process. The problem associated with the estimation of parameters of phase type distribution is solved by the EMpht algorithm. The methodology suggested can be applied in various similar situations in seed physiology and more generally to problems of similar nature in Plant science.

6.5 Tables and Figures:

Table - I : Data on the time taken (in hrs.) by 100 seeds in the four growth stages.

seed	StageI	StageII	StageIII	StageIV	seed	StageI	StageII	StageIII	Stage IV
1	27	26	21	45	51	29	26	29*	–
2	29	31	25	46	52	25	28	11^v	–
3	32	29	26	15^v	53	29	31	26	54
4	26	29	24	44	54	28	34	25	9^v
5	31	33	9^v	–	55	35*	–	–	–
6	23	24	22	55*	56	35*	–	–	–
7	35*	–	–	–	57	32	36	29*	–
8	30	37*	–	–	58	27	28	12^v	–
9	31	29	21	42	59	27	28	12^v	–
10	14	29	24	45	60	35*	–	–	–
11	32	35	18^v	–	61	33	37*	–	–
12	26	25	23	55*	62	34	37*	–	–
13	35*	–	–	–	63	28	30	6^v	–
14	33	35	29*	–	64	27	26	29*	–
15	28	30	24	23^v	65	27	10^v	–	–
16	25	24	20	44	66	28	37*	–	–
17	25	25	9^v	–	67	25	25	22	41
18	27	31^v	–	–	68	25	25	29*	–
19	28	20^v	–	–	69	29	24^v	–	–
20	32	34	28	51	70	35*	–	–	–
21	34	32	23	45	71	28	37*	–	–
22	24	26	20	40	72	35*	–	–	–
23	31	33	12^v	–	73	27	32	26	47
24	29	34	8^v	–	74	26	29	20^v	–
25	29	30	26	55*	75	33	19^v	–	–
26	30	34	14^v	–	76	35*	–	–	–
27	35*	–	–	–	77	31	27	21	43
28	25	37*	–	–	78	29	33	10^v	–
29	31	37*	–	–	79	35*	–	–	–
30	35*	–	–	–	80	29	37*	–	–
31	26	24	29*	–	81	30	37*	–	–
32	25	29	26	44	82	27	27	24	55*
33	33	31	25	48	83	25	13^v	–	–
34	31	36	10^v	–	84	29	32	4^v	–
35	34	31	29*	–	85	32	37*	–	–
36	35*	–	–	–	86	26	18^v	–	–
37	35*	–	–	–	87	35*	–	–	–
38	25	32	28	18^v	88	24	26	23	47
39	33	36	29*	–	89	29	33	27	53
40	26	27	29	22^v	90	35*	–	–	–
41	31	30	28	55*	91	30	37*	–	–
42	34	32	28	50	92	35*	–	–	–
43	26	27	29	22^v	93	26	14^v	–	–
44	28	33	7^v	–	94	35*	–	–	–
45	27	29	29*	–45	95	32	33	28	52
46	29	24^v	–	–	96	33	32	27	55*
47	35*	–	–	–	97	35*	–	–	–
48	35*	–	–	–	98	35*	–	–	–
49	32	37*	–	–	99	29	30	25	47
50	27	37*	–	–	100	35*	–	–	–

* indicates the censoring times and ^{v} indicates the decay times

Table - II

Observed decay/censoring times (in hrs.) of the 100 seeds.

(^{v} denotes the decay times)

seed	Decay/censoring time	seed	Decay/censoring time
1	119	51	84
2	131	52	64^v
3	102^v	53	140
4	123	54	96^v
5	73^v	55	35
6	124	56	35
7	35	57	97
8	67	58	67^v
9	123	59	84
10	112	60	35
11	85^v	61	70
12	129	62	71
13	35	63	64^v
14	97^v	64	82
15	105	65	37^v
16	113	66	65
17	59^v	67	113
18	58	68	89
19	48	69	53^v
20	145	70	35
21	144	71	65
22	110	72	35
23	76^v	73	132
24	71^v	74	75^v
25	140	75	52^v
26	78^v	76	35
27	35	77	122
28	62	78	72^v
29	68	79	35
30	35	80	66
31	79	81	67
32	124	82	133
33	137	83	38^v
34	77^v	84	65^v
35	94	85	69
36	35	86	44^v
37	35	87	35
38	103^v	88	120
39	98	89	142
40	104^v	90	35
41	144	91	67
42	144	92	35
43	104^v	93	40^v
44	68^v	94	145
45	85	95	35
46	53^v	96	147
47	35	97	35
48	35	98	35
49	69	99	131
50	64	100	35

FIGURE - 1

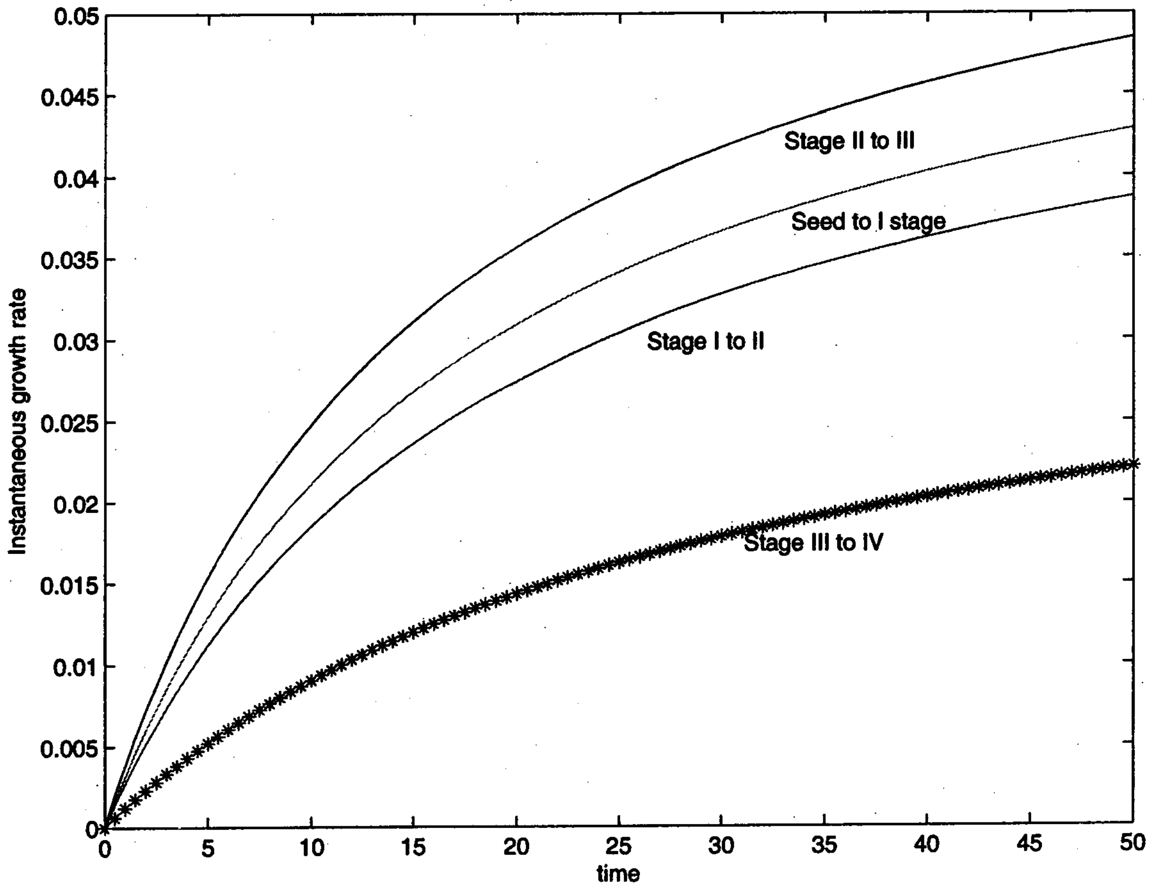


FIGURE - 2

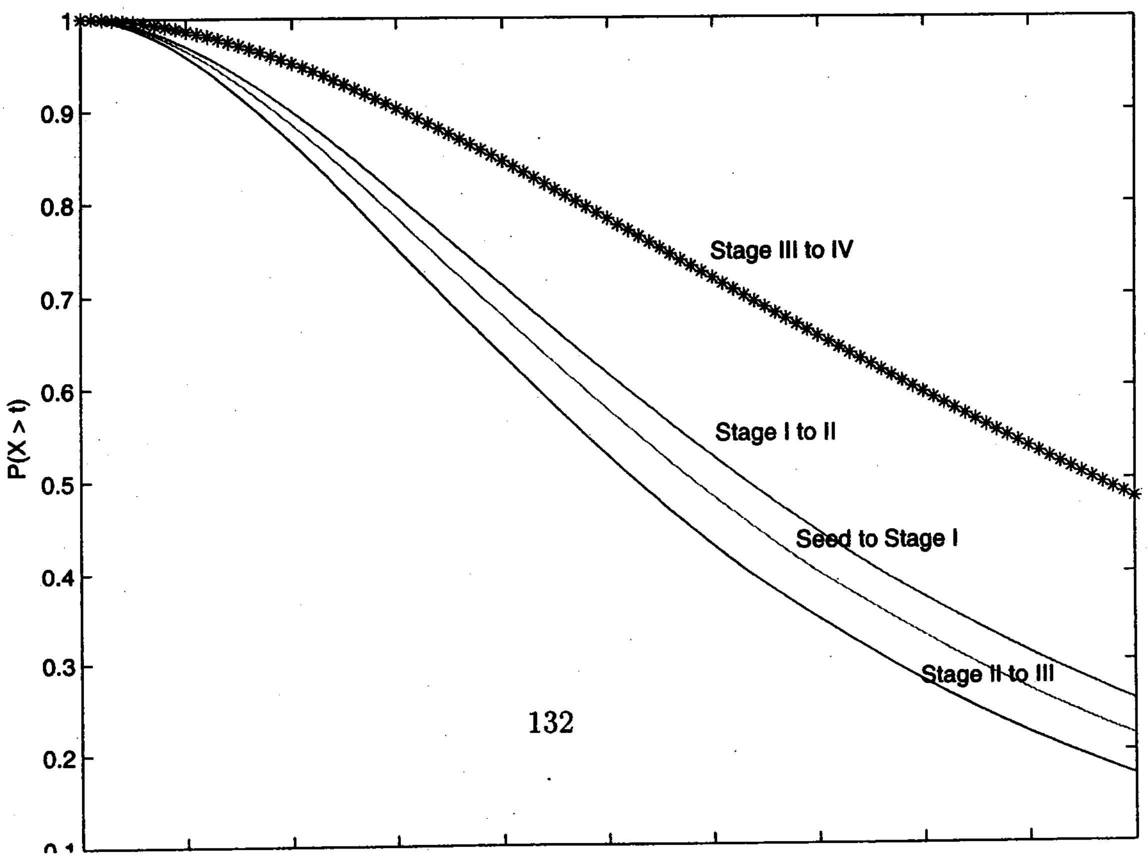


FIGURE - 3

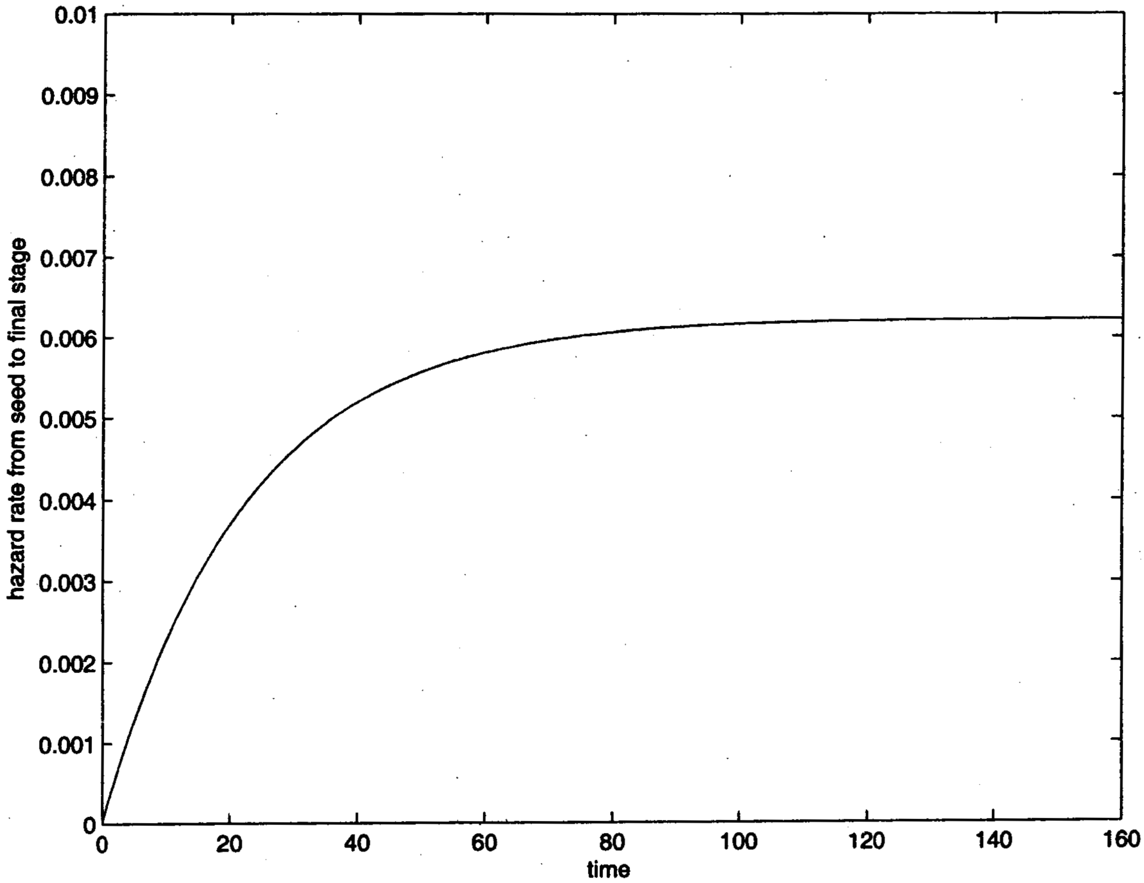
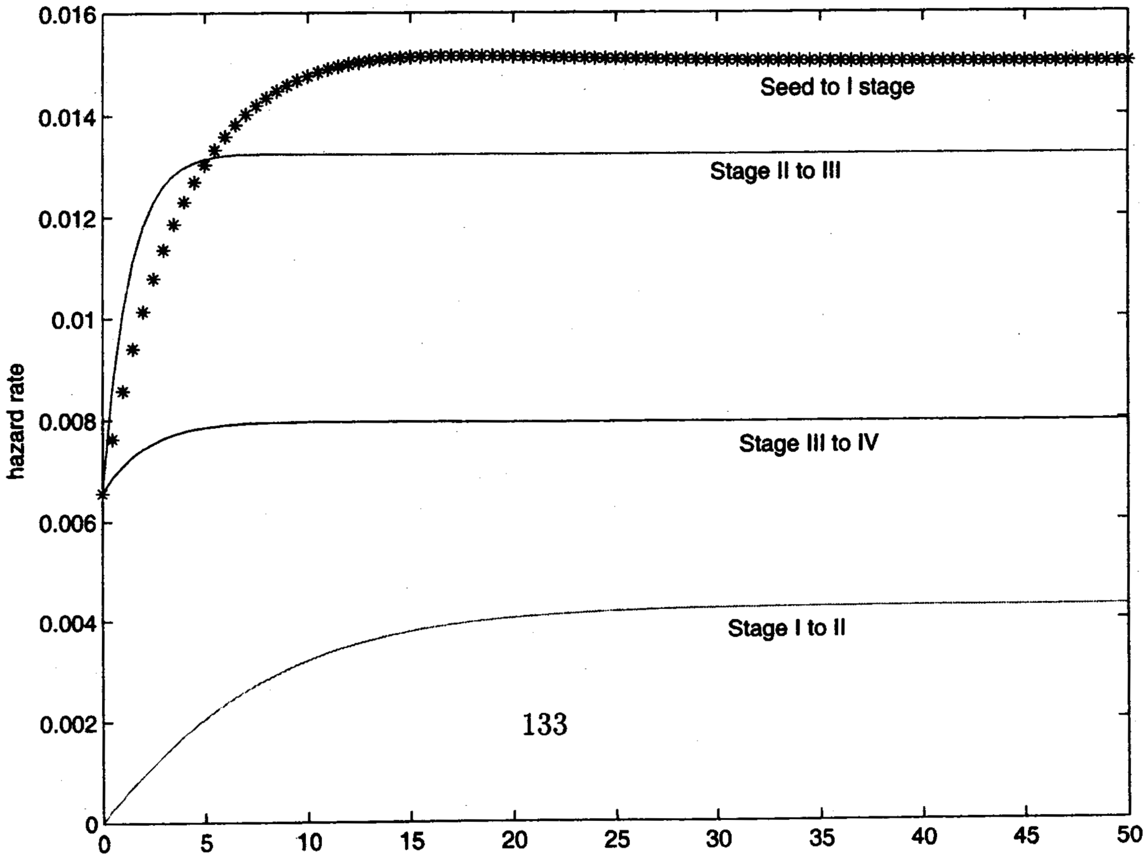


FIGURE 4



Concluding Remarks

Phase type distribution has been introduced as a tool for unifying a variety of stochastic models and found useful initially in many queuing models. Its application in reliability and survival analysis, though initiated sometime ago, is found to be quite promising now. There is a recent surge of interest in applying phase type models in industrial engineering and biomedical sciences. The present thesis augments a great deal towards the applications of this important probability distribution in reliability and survival analysis.

The characterization of phase type distribution has been attempted by many researchers in the past. In *chapter 2*, we presented a new characterization result of phase type distribution under Poisson shocks. Characterization or identifiability of the shock models based on the life distribution of the system is an interesting research problem one can investigate further in this regard.

Chapter 3 contributes a solution for a reliability and warranty problem under general lifetime distribution assumption. Using the method illustrated many problems in business management under the assumption of general lifetime distribution for a device/product, can be solved effectively. The methodology suggests a practical solution to the warranty problems of products under random environment. Also it is worth investigating the role of various warranty policies when the underlying life distribution is of phase type.

A problem of preventive replacement of a device under Poisson shock is considered in the *chapter 4*. The problem assumes greater significance when loss due to preventive replacement is smaller than that due to the corrective replacement. For discrete phase type failure probabilities, the problem of finding the optimum value of the replacement time by minimising long run cost per unit time for replacement, and the same problem under a given loss function are discussed. A comparison of the above preventive replacement policies with the failure replacement is made proposing the gain due to age replacement.

In *chapter 5*, a multistate system under Poisson shock is considered. While the lifetime of the system is distributed as continuous phase type, the reliability and hazard plots of the system are obtained. An extension of the lifetime modelling of a multistate system using phase type distribution is possible in the field of medicine. The lifetime of a human being under the threat of heart attack or cerebral hemorrhage can be modelled by extending the method presented in this chapter.

In the last chapter, a problem of modelling seed germination under general life distribution assumption using phase type distribution is discussed. We considered different growth stages of the seed and the survival, hazard rate of the seed on its the different growth stages are obtained. The method discussed is capable for helping plant scientists. A more close observation on seed germination of a given species can be made by considering more stages on germination and the effect of random environment on seed germination also can be incorporated.

Appendix

Computer Codes

B.1 Chapter 3

No.1 ***[To plot reliability and hazard function of the component]***

```
T=[-.6 .1 .2;.5 -1.7 .1;.1 .4 -1.6]
B=0.5;
j-1;
a=[1 0 0]
for i=1:2;
i(j)=I;
p(j)=b^i(j);
j=j+1;
end
G=[t,eye(3,6)*0;t*0,T*P(1),eye(3,3)*0;eye(3,6)*0,T*p(2)]
e=[.7 .1 .2];
a1=Kron(e,a);
f=Kron(G,eye(9^2))+Kron(G,eye(9))+Kron(eye(9^2),G)
g11=g1';b1=(-1)*(sum(g(1));b=b1';
q=1;
m=1;
for t=0:.2:1;
tt(m)=t;
s(q)=f1*(expm(g1*t))*b;m=m+1;
q=q+1;
end
e1=sum(Kron(eye(9),eye(9^2)));e=e1';r=1;
for k=0:.2:1
kk(r)=k;
Q(r)=(f1*(expm(g1*k))*b)/(f1*(expm(g1*k)))*e;
r=r+1;
end
tt
s
Q
Kk
Plot(kk,Q)
```

No.2 ***[To decide the number of preventive maintenances]***

```
T=[-.6,.1,.2;.5,-1.7,.1;.1,.4,-1.6]
a=[.7,0,0,.1,0,0,.2,0,0]
e=[1,1,1,1,1,1,1,1,1]'
L=[T,eye(3)*0,eye(3)*0;eye(3)*0,T/.8,eye(3)*0;eye(3)*0,eye(3)*0,T/.5]
L1=[T/.8,eye(3)*0,eye(3)*0;eye(3)*0,T/(.8*.8),eye(3)*0;eye(3)*0,eye(3)*0,T/(.8*.5)]
L2=[T/.6,eye(3)*0,eye(3)*0;eye(3)*0,T/(.6*.8),eye(3)*0;eye(3)*0,eye(3)*0,T/(.6*.5)]
F=(a*expm(L*.5)*e)
Q1= -log(F)
F1=(a*expm(L1*.4)*e)
Q2= -log(F1)
F2=(a*expm(L2*.34)*e)
Q3= -log(F2)
C1=10000*Q1+150
C2=10000*(Q1+Q2)+300
C3=10000*(Q1+Q2+Q3)+450
```

B.2 Chapter 4

No.1[*To plot the loss function*]

```
T=[-.7 .2;.3 -.4];
a=[1,0];
e=[1;1];
c1=1;
c2=7;
L=-a*(inv(T))*e
j=1;
for t=0:.4:L
   k(j)=t;
   kk(j)=c1*(L-t)^2+c2*(-log(a*expm(T*t)*e));
   j=j+1;
end
k
kk
plot(k,kk)
```

No.2[*To find the gain due to preventive replacement*]

```
T=[(-.6),.2;.4,-.4]
a=[.6,.4];
e=[1;1];
c1=10;
c2=8;
L=-a*(inv(T))*e
k=c1*(L-2.25)^2+c2*(-log(a*expm(T*(2.25))*e))
V=50
M=5000
aa=(a*expm(2.25*T)*e)
bb=(1-(a*expm(2.25*T)*e))
Caa=[(V+k)*(a*expm(2.25*T)*e)]+[(M)*(1-(a*expm(2.25*T)*e))]
Cbb=[(a*expm(2.25*T)*inv(T)*e)-a*inv(T)*e]
Ca=Caa/Cbb
Pi=(1/L)*a*inv(-T)
P=[0.5263,0.4737;0.5263,0.4737]
pp=50000
Ren=(L)^(-1)*pp-1+((a*T^(-2)*e)/(L)^2)+(L)^(-1)*a*(P-expm(T+(e-T*e)*a)*pp)*inv(T)*e
Cf=M*Ren/50000
Gain=Cf-Ca
```

B.3 Chapter 5

No.1[*To plot the reliability and hazard function*]

```
T=[0,.634,0.238,0.089,0.028;0,0,.634,.238,.089;0,0,0,.634,0.238;0,0,0,0,.634;0,0,0,0,0]
a=[1,0,0,0,0]
e=[1;1;1;1;1];
l=2;
I=eye(5)
b=[.011;.039;.128;.366;1];
L=-a*(inv(l*(T-I)))*e
j=1;

for t=0:.4:3*L
   k(j)=t;
   kk(j)=(a*expm(l*(T-I)*t)*b)/(a*expm(l*(T-I)*t)*e);
   kkk(j)=a*expm(l*(T-I)*t)*e;
   j=j+1;
end
k
kk
kkk
plot(k,kk,k,kkk)
```

No.2[*To plot the reliability and hazard function*

```
T=[0,.3975,0.2387,0.1447,0.0878;0,0,.3975,.2387,.1447;0,0,0,.3975,0.2387;0,0,0,0,.3975;0,0,0,0,0]
a=[1,0,0,0,0]
e=[1;1;1;1;1];
l=2;
I=eye(5)
b=[.1353;.2191;.3638;.6025;1];
L=-a*(inv(l*(T-I)))*e
j=1;

for t=0:.4:3*L
   k(j)=t;
   kk(j)=(a*expm(l*(T-I)*t)*b)/(a*expm(l*(T-I)*t)*e);
   kkk(j)=a*expm(l*(T-I)*t)*e;
   j=j+1;
end
k
kk
kkk
plot(k,kk,k,kkk)
```

B.4 Chapter 6

No.1[*To plot instantaneous growth rate of the seed in each its different growth stages*]

```
a=[1,0]
b=[1,0]
c=[1,0]
d=[1,0]
k=[-0.057735,0.057735;0,-0.057735]
l=[-0.053238,0.053238;0,-.053238]
m=[-0.06366,0.06366;0,-0.06366]
n=[-0.034885,0.034885;0,-0.034885]
kk=-1*sum(k')
e=[1;1]
j=1;
for t=0:.5:50;
  h(j)=t;
  hh(j)=(a*expm(k*t)*kk')/(a*expm(k*t)*e)
  j=j+1;
end
h
hh
ll=-1*sum(l')
j=1;
for t=0:.5:50;
  p(j)=t;
  pp(j)=(b*expm(l*t)*ll')/(b*expm(l*t)*e)
  j=j+1;
end
p
pp
mm=-1*sum(m')
j=1;
for t=0:.5:50;
  q(j)=t;
  qq(j)=(c*expm(m*t)*mm')/(c*expm(m*t)*e)
  j=j+1;
end
q
qq
nn=-1*sum(n')
j=1;
for t=0:.5:50;
  r(j)=t;
  rr(j)=(d*expm(n*t)*nn')/(d*expm(n*t)*e)
  j=j+1;
end
r
rr
plot(h,hh,'-g')
hold on
plot(p,pp,'-r')
hold on
plot(q,qq,'-b')
hold on
plot(r,rr,'*r')
```

No.2[*To plot the survival function of the seed in various stages*]

```
a=[1,0]
b=[1,0]
c=[1,0]
d=[1,0]
k=[-0.057735,0.057735;0,-0.057735]
l=[-0.053238,0.053238;0,-.053238]
m=[-0.06366,0.06366;0,-0.06366]
n=[-0.034885,0.034885;0,-0.034885]
kk=-1*sum(k')
e=[1;1]
j=1;
for t=0:.5:50;
  h(j)=t;
  hh(j)=(a*expm(k*t)*e)
  j=j+1;
end
h
hh
ll=-1*sum(l')
j=1;
for t=0:.5:50;
  p(j)=t;
  pp(j)=(b*expm(l*t)*e)
  j=j+1;
end
p
pp
mm=-1*sum(m')
j=1;
for t=0:.5:50;
  q(j)=t;
  qq(j)=(c*expm(m*t)*e)
  j=j+1;
end
q
qq
nn=-1*sum(n')
j=1;
for t=0:.5:50;
  r(j)=t;
  rr(j)=(d*expm(n*t)*e)
  j=j+1;
end
r
rr
plot(h,hh,'-g')
hold on
plot(p,pp,'-r')
hold on
plot(q,qq,'-b')
hold on
plot(r,rr,'*r')
```

N0.3[To ***plot the hazard rate of the seed in various stages of growth***]

```
a=[0.001334,0.800287,0.198378]
k=[-0.084882,0.017825,0.05975;0.046926,-0.166635,0.119709;0.151992,0.051021,-0.203013]
kk=-1*sum(k')
e=[1;1;1]
j=1;
for t=0:.5:50;
   h(j)=t;
   hh(j)=(a*expm(k*t)*kk')/(a*expm(k*t)*e)
   j=j+1;
end
h
hh
b=[0.804312,0.195688,0.000001]
l=[-0.358696,0.212524,0.146172;0.240711,-0.507013,0.266303;0.128051,0.150568,-0.311638]
ll=-1*sum(l')
e=[1;1;1]
j=1;
for t=0:.5:50;
   h2(j)=t;
   hh2(j)=(a*expm(l*t)*ll')/(a*expm(l*t)*e)
   j=j+1;
end
h2
hh2
c=[0.203247,0.394742,0.402011]
m=[-0.253126,0.166938,0.083404;0.105748,-0.205621,0.096256;0.115563,0.161631,-0.292831]
mm=-1*sum(l')
e=[1;1;1]
j=1;
for t=0:.5:50;
   h3(j)=t;
   hh3(j)=(a*expm(m*t)*mm')/(a*expm(m*t)*e)
   j=j+1;
end
h3
hh3
d=[0.999965,0.000035,0]
m1=[-0.167049,0.167044,0.000005;0.017308,-0.131828,0.11452;0.027979,0.058696,-0.11937]
mm1=-1*sum(l')
e=[1;1;1]
j=1;
for t=0:.5:50;
   h4(j)=t;
   hh4(j)=(a*expm(m1*t)*mm1')/(a*expm(m1*t)*e)
   j=j+1;
end
h4
hh4
plot(h,hh,'-g')
hold on
plot(h2,hh2,'-r')
hold on
plot(h3,hh3,'-b')
plot(h4,hh4,'*b')
```

No.4 To ***plot the hazard rate of the seed from its first stage to final stage of growth***]

```
a=[0.312411,0.687589,0]
k=[-0.032552,0.004845,0.027707;0.044172,-0.065707,0.021535;0.006152,0.012213,-0.030373]
kk=-1*sum(k')
e=[1;1;1]
j=1;
for t=0:1:200;
   h(j)=t;
   hh(j)=(a*expm(k*t)*kk')/(a*expm(k*t)*e)
   j=j+1;
end
h
hh
plot(h,hh)
axis([0 160 0 .01])
```

Bibliography

Bibliography

[1] Aalen, O.O. (1995). Phase type distribution in survival analysis, *Scand. J. Statist.*, **2**, 447-463.

[2] Ahlstrom, L., Olsson,M. and Nerman,O.(1999), A parametric estimation procedure for relapse time distribution, *Lifetime Data Analysis,* **5**, 113–132.

[3] Aldous, D. and Shepp, L. (1987), The least variable phase-type distribution is Erlang, *Commun. Statist. Stochastic Models*, **3**, 467-473.

[4] Asmussen, S. (1992), Phase-type distributions in random walk and queuing problems, *Ann. Probab.*, **20**: 772-789.

[5] Asmussen, S.,Nerman, O. and Olsson, M. (1996), Fitting Phase-type distributions via the EM algorithm, *Scand. J. Statist.*, **23**: 419-441.

[6] Asmussen, S. and Rolski, T. (1991), Computational methods in Risk theory : a matrix-algorithmic approach, *Insurance Math. Econom.*, **10**: 259-274.

[7] Assaf, D.,Langberg, N.A.,Savits, T.H. and Shaked, M. (1984), Multi variate phase type distribution, *Oper. Res.*, **32**, 688-701.

[8] Assaf, D. and Levinskon, B. (1982), Closure of phase type distribution under operations arising in reliability theory, *Ann. Prob.*, **10**, 265-269.

[9] Assaf, D.,Naftali, A. and Langberg. (1985), Presentation of phase type distributions as proper mixtures, *J. Appl. Prob.*, **22**, 247-250.

[10] Aven, T. and Jensen, U. (1999), *Stochastic Models in Reliability*, Springer Verlag, New York.

[11] Barlow, R.E. and Proschan, F. (1965), *Mathematical Theory of Reliability*, John Wiley, New York.

[12] Barlow, R.E. and Proschan, F. (1975), *Statistical Theory of Reliability and Life testing*, Holt, Rinehart and Winston Inc, New York.

[13] Barlow, R.E. and Wu, A.S.(1978), Coherent system with multistate components, *Mathematics of Operations Research.* **4**, 275-281.

[14] Baron, Y., Frostig, E. and Levikson, B. (2006), Analysis of R out of N systems with several repairmen, Exponential life times and phase type repair times : An algorithmic approach, *European Journal of Operational Research* **169**, 1, 202-225.

[15] Berenguer, C.,Grall, A. and Dieulle,L. (2003), Maintenance policy for a continuously monitored deteriorating system, *Probability in Engineering and Informational Sciences.* **17**, 235-250.

[16] Birnbaum, Z.W.; Esary, J.D. and Saunders, S.C. (1961), Multi component systems and structures and their reliability, *Technometrics*, **3**, 55-77.

[17] Bobbio, A and Telek, M. (1994), A bench mark for PH estimation algorithms : Results for acyclic-PH , *Commun. Statist.-Stochastic Models*, **10(3)**, 661-667.

[18] Bobbio, A., Horvath, A., Scarpa, M. and Telek, M. (2003), Acyclic discrete phase type distributions : properties and a parameter estimation algorithm, *Performance Evaluation*, **54**, 1-32.

[19] Bobbio, A.,Horvath, A. and Telek, M. (2005), Matching three moments with minimal acyclic phase type distributions, *Stochastic Models*, **21**, 303-326.

[20] Bryson, M.C. and Siddiqui, M.M. (1969). Some criteria for ageing, *J. Amer. Statist. Ass.*, **64**, 1472-1483.

[21] Chakravarthy, S.(1983), Reliability analysis of a parallel system with exponential lifetimes and phase type repairs, *O R Spectrum* **5**, 25-32.

[22] Chakravarthy, S.(1987), Analysis of production line systems with two unreliable machines with phase type processing times and a finite storage buffer, *Commns. in Stats : Stochastic Models,* **3**, 369–391.

[23] Chakravarthy, S.R., Krishnamoorthy, A. and Ushakumari, P.V(2001), A k-out-of-n reliability system with an unreliable server and phase type repair and services : The (N,T) policy, *Journal of Applied Mathematics and Stochastic Analysis* **14:4**, 361-380.

[24] Chen, T. and Popova, E. (2002), Maintenance Policies with two-dimensional warranty.*Reliability Engineering and System Safety.* **77**, 61-69.

[25] Cho, D.I. and Parlar, M. (1991), A survey of maintenance models for multi-unit systems, *European Journal of Operations Research,* **51**, 01-23.

[26] Copeland, L.O., (1988), *Introduction to Seed Technology,* Surjeet publications, New Delhi.

[27] Cox, D.R. (1955),The analysis of non-Markovian Stochastic processes by the inclusion of supplementary variables, Louvain,Belgium,173-206.

[28] Cox.D.R. (1972), Regression models and life tables (with discussion), *Journal of Royal Statistical Society, Series B* **34**, 187-220.

[29] Cui, L. and Li, H.(2007), Analytical method for reliability and MTTF assessment of coherent systems with dependent components, *Reliability Engineering and System Safety,* **92(3)**, 300–307.

[30] Day, N.E. and Brown, C.C. (1980), Multistage models and primary prevention of cancer. *J. Natl Cancer Inst.* **64**, 977-989.

[31] Dempster, A.P.,Laird, N.M. and Rubin, D.B. (1977), Maximum likelyhood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc. Ser. B* **39**, 1-38.

[32] Dharmadhikari, A.D. and Manoharan, M. (1995), On a closure problem in multistate monotone system, *Operations Research: Theory and Practice,* Ed: P.K. Kapur and V.K. Sehgal, Spaniel Publications, New Delhi, 63-70.

[33] El-Neweihi,E.,Proschan,F.,and Sethuraman,J. (1978), Multistate coherent systems, *J. Appl.Prob.*, **15**,675-688.

[34] Epstein, B. and Sobel, M. (1953), Life testing, *J. Amer. Statist. Ass.*, **48**, 486-502.

[35] Faddy, M.J. and McClean, S.I. (2000), Analysing data on lengths of stay of hospital patients using phase type distribution, *Applied Stochastic Models in Business and Industry* **15(4)**, 311-317.

[36] Grifith, W.S., (1980), Multistate reliability models, *J. Appl. Prob.* **17**, 735-744.

[37] Gururajan, M. and Bhat, K.S. (1990), A complex priority redundant system with phase type distribution, *Microelectronics and Reliability* **30**, 453-455.

[38] Hardy, G.H. Littlewood, J.E. and Polya, G. (1952), *Inequalities*, 2nd edition, Cambridge Univ. Press.

[39] Horvath, A. and Telek, M. (2002), Phfit : A general Phase-type fitting tool, *Proc. of 12th performance TOOLS, Lecture notes in Computer Science 2324*.

[40] Huggstrom, S.,Asmussen, S. and Nerman, O. (1992), EMPHT - A program for fitting phase type distributions, *Technical Report*, Dept. of Maths.,Chalmers Uty. of Tech.;Goteborg, Swedon.

[41] Jenson, A. (1954), *A distribution model applicable to economic*, Munksgard, Copenhagen.

[42] Jonson, E., Anderson, M. and Asmussen, S. (1994), A practical dependability measure for degradable computer system with non exponential degradation. *Proc. of Safe Process '94* , Helsinky, June 1994.

[43] Jack, N. and Dagpunar, J.S (1994), Preventive Maintenance strategy for equipment under warranty, *Microelectronics and Reliability.* **34**, 1089-1093.

[44] Jack, N. and Murthy, D.N.P. (2002), A new preventive maintenance strategy for item sold under warranty, *IMA Journal of Management Mathematics.* **13**, 121-129.

[45] Kao, E.P.C. (1988), Computing the phase type renewal and related functions, *Technometrics*, **15**,675-688.

[46] Kaplan, E.L. and Meier, P. (1958), Non parametric estimator from incomplete observations, *Journal of American Statistical Association* **53**, 457-481.

[47] Katayama, T., Okamoto, M. and Enamoto, H. (1978), Characterization of the structure-generating functions of regular sets and the DOL growth functions,*Inform. and Control.*, **36**, 85-101.

[48] Keilson, J. (1975), Monotonicity and convexity in system survival functions and Metabolic disappearance curves, in reliability and biometry, *Statistical Analysis of Life Lenghts*, 81-98, SIAM, Philadelphia.

[49] Latouche, G. and Ramaswami, V. (1999), *Introduction to Matrix Analytic Methods in Stochastic Modelling*, SIAM & ASA, Philadelphia.

[50] Lawless, J.F. (1982), *Statistical Models and Methods for Lifetime Data,* Wiley, New York.

[51] Lipsky, L. (1992), *Queuing Theory : a Linear Algebraic Approach*, Mac Millan, New York, B,D.

[52] Maier, R.S., and O'Cinniede, C.A. (1992), A closure characterization of phase type distribution, *Journal of Applied Probability*, **29(1)**, 92–103.

[53] Manoharan, M., Harshinder Singh and Neeraj Misra (1992), Preservation of phase type distribution under Poisson shock models, *Adv. Appl. Prob.*, **24**, 92-103.

[54] Manoharan, M. and Joseph, K.X.(2004), A multistate deteriorating system and its optimal replacement schedule, *Journal of Kerala statistical association.* **15**, 01–12.

[55] Mantel, N. (1966), Evaluation of survival data and new order statistics assuming in its consideration, *Cancer Chemotherapy Report* **50**, 163-170.

[56] Marshal, A.H. and Mc Clean, S.I. (2003), Conditional phase type distribution for modelling patient lengths of stay in hospital, *International Transactions in Operations Research* **10(6)**, 565-576.

[57] Marshall, A.W. and Olkin, I. (1979), *Inequalities: The theory of Majorization and its Applications.* Academic Press, New York.

[58] Michael E. O'Neill, Peter C. Thomson, Brent C. Jacobs, Phil Brain, Ruth C. Butler, Heather Turner and Bernadetha Mitakda (2004), Fitting and comparing seed germination models with a focus on the inverse normal

distribution, *Australian & New Zealand Journal of Statistics,* **46(3)**, 349–366.

[59] Murthy, D.N.P. and Blischke,W.R. (2001), Warranty and Reliability, Chapter 20, HandBook of Statistics,Vol.20,Advances in Reliability, 541-583.

[60] Murthy, D.N.P., Iskandar, B.P. and Wilson, R.J. (1990), Two dimensional failure free warranties; Two-dimensional point process models, *Operations Research.* **43**, 356-366.

[61] Narasimhalu, Y.C. and Prasad, B.S.(1997), The accumulated damage model-Optimum replacement policy, *IAPQR Transactions*, **22**, 11-21.

[62] Neuts, M.F.(1975), *Probability distributions of phase type*, Liber Amicorum Professor Emeritus H.Florin, University of Louvain,Belgium.,173-206.

[63] Neuts, M.F. (1981), *Matrix-Geometric Solutions in Stochastic Models: An Algorithm approach*, Johns Hopkins University Press, Baltimore.

[64] Neuts, M.F.(1989), *Structured stochastic matrices of M/G/1 type and their applications*, Marcel Dekker, NewYork.

[65] Neuts, M.F. (1992), Two further closure properties of PH-distributions, *Asia-Pacific J. Oper. Res.*, **9**, no.1, 77-85.

[66] Neuts, M.F. and Meier, K.S.(1981), On the use of phase type distributions in reliability modelling of system with a small number of components. *O.R.Spektrum.* **2**, 227-234.

[67] Ngüygen, D.G. and Murthy, D.N.P.(1989), Optimal replace-repair strategy for servicing products sold with warranty. *European Journal of Operational Reasearch.* **39**, 206-212.

[68] Nguyen, D.G. and Murthy, D.N.P.(1986), An optimal policy for servicing warranty. *Journal of Operational Research Society.* **11**, 1081-1088.

[69] O'Cinneide, C.A. (1990), Characterization of phase-type distributions, *Commun. Statist. Stochastic Models.* **6**, 1-57.

[70] O'Cinneide, C.A. (1991), Phase type distributions and majorization, *Ann. Appl. Probab.* **1**, no.2, 219-227.

[71] O'Cinneide, C.A. (1999), Phase type distributions : Open problems and a few properties, *Commun. Statist. Stochastic Models.* **15**, 731-757.,

[72] Olsson, M. (1996), Estimation of Phase type distributions from censored data, *Scand. J. Statist.* **23**: 443-460.

[73] Olsson, M. (1998), The EMpht program, *Technical Report* , Dept. of Math. Chalmers Uty. of Tech.;Goteborg, Swedon.

[74] Renhai, Y. (2002), A proof of the steepest increase conjecture of a phase type density, *Stochastic Models,* **18(1)**,1–6.

[75] Sengupta, B. (1989), Markov process whose steady state distribution is matrix exponential with an application to the G1/G/1 queue, *Adv. Appl. Prob..*, **21**, 159-180.

[76] Shafili, B. and Price, W.J.(2001), Estimation of cardinal temperatures in germination data analysis, *Journal of Agricultural, Biological & Environmental Statistics,* **6(3)**, 356–366.

[77] Shantikumar, J.G.(1985), Bilateral phase type distributions, *Naval Research Logistic* **32**, 119–136.

[78] Shi Dinghua., Guo Jinli and Liu Liming. (2005), On SPH Distribution class, *Acta Mathematica Scientia,* **25(2)**,201–214.

[79] Shi Dinghua, Guo, J.L. and Liu, L.M., (1996), SPH-distributions and the rectangle-iterative algorithm, *In : Chakravarthy, Alfa eds. Matrix Analytic Methods in Stochastic Model. New York : Marcel Dekker*, 207–224.

[80] Soittola, M. (1976), Positive rational sequences, *Theoretical Computer Science,* **2**, 317–322.

[81] Soohan Ahn and Ramaswami,V. (2005), Bilateral phase type distributions, *Stochastic Models,* **21**, 239–259.

[82] Suprasad Amari and Wes Fulton,(2003), Bounds on optimal replacement time of age replacement policy, *Proceedings of Annual Reliability and maintanability symposium, Tampa,Florida,27-30Jan2003*, 417-422.

[83] Telek, M. (July 2000), The minimal coefficient of variation of DPH distributions, *Proceedings of the 3rd international conference on matrix-analytic methods in stoch.models,*.

[84] Wei Li and Alfa,A.S. (2005), Scheduling stochastic jobs on a repairable machine with general phase type uptime, *Mathematical Methods of Operations Research,* **61**,399–417.

[85] Wu, C.F.J., (1983), On the convergence properties of the EM algorithm, *Ann. Statist.* **11**, 95-103.