University of Nairobi

School of Mathematics

ISSN: 2410-1397

Master Project in Actuarial Science

Using Phase-Type Distribution to Determine Actuarial Functions for Whole Life Insurance Policies

Research Report in Mathematics, Number 29, 2020

Olondo Utshudi Solange

November 2020

Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Actuarial Science

Using Phase-Type Distribution to Determine Actuarial Functions for Whole Life Insurance Policies

Research Report in Mathematics, Number 29, 2020

Olondo Utshudi Solange

School of Mathematics
College of Biological and Physical sciences
Chiromo, off Riverside Drive
30197-00100 Nairobi, Kenya

Master Thesis

Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Actuarial Science

Submitted to: The Graduate School, University of Nairobi, Kenya

Abstract

Background The world life expectancy has had a continual increase for both male and female gender. This increasing trend is expected to continue, hence, the need to have a model that does not have a maximum attainable age. The Lee-carter model suggests that there is no limiting age, however, this model is not sufficient for computation of actuarial functions . On the other hand, the De Moivre model which was purposely built for computation of annuity functions, and possibly other actuarial functions, requires a limiting age. There is therefore need to price life contracts by not constraining the future life time of policy holders to a maximum attainable age.

Methodology This study applies the phase-type model on AM92 data to compute actuarial functions for whole life insurance policies not limiting the future lifetime of individuals to a maximum attainable age. Assurances and annuities were determined by induction of the Laplace function, moment and probability generating functions of the phase-type distribution.

Results and Conclusion The phase-type distribution of Coxian nature is applicable for computation of functions. The one absorbing state is death and ages form the phases of the distribution. A life in state (i) transits to the next phase with rate λ while it dies with rate $q_i = a + bi^c$, where, a is death parameter due to accident while bi^c is due to biological aging of the life. The parameters λ, a, b and c were estimated from the Nelson-Mead algorithm in that the aim was to minimize the mean squared error of the survival function. Results showed that there is small margin of deviation between premiums computed directly from life table functions and those computed by applying a Phase-Type distribution. Also, there is a large underestimation of assurances and overestimation of annuities between the ages 40 and 90. However, there are small deviations for whole life net premiums.

Master Thesis in Mathematics at the University of Nairobi, Kenya.
ISSN 2410-1397: Research Report in Mathematics

DISTRIBUTOR: School of Mathematics, University of Nairobi, Kenya

Declaration and Approval

I the undersigned declare that this dissertation is my original work and to the best of my knowledge, it has not been submitted in support of an award of a degree in any other university or institution of learning.

Signature Date

Olondo Utshudi Solange

Reg No. I56/14213/2018

In my capacity as a supervisor of the candidate's dissertation, I certify that this dissertation has my approval for submission.

Signature Date

Prof JAM Otieno
School of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: jam@uonbi.ac.ke

Dedication

I dedicate this project to the almighty God for his protection and providence. My strengths, health, knowledge and understanding are sourced from Him. This work is also dedicated to my mother and my entire family for continuing to shower me with love and care. Thank you for your unending emotional, spiritual and financial support. You are indeed a big part of my success. Thank you for trusting in me. I also dedicate this project to my church, Renewed Pentecostal Missionary Church for their prayers and support.

Contents

Acknowledgments

First of all, I thank the all mighty God for this grace, providence and unending love. He has provided for my every need and I never lacked anything that was necessary for me to complete this course. I cannot overstate my deepest and heartfelt gratitude to my loving family, most specially my mother Mrs. Utshudi, for their continued spiritual, financial and emotional support. My heartfelt gratitude are also extended to Renewed Pentecostal Missionary Church for their prayers and support.

I would like to give my deepest appreciation to my supervisor Prof. JAM Ottieno for his unending support, guidance, and trust all through this course. I also appreciate Prof. POG Weke for making sure I joined this course and completed my project. I would also like to extend my gratitude to the University of Nairobi, particularly, the School of Mathematics for the opportunity to complete this course. Last but not least I would like to thank all my lectures, colleagues and friends for their continued support all through the years.

Olondo Utshudi Solange

Nairobi, 2020.

1 Introduction

1.1 General Introduction

Insurance policies are contracts between an insurance company and a policy holder, where the former receives an amount called premium from the latter in exchange to cover a risk of the insured life or property. Actuaries set premiums by incorporating uncertainties about the time when the insured risk would occur, the magnitude of the claim, and the time value of money. In order to make profits, insurance companies set premiums such that the mean present value of premiums and expenses supersedes that of benefits.

Traditional insurance contracts include life assurance policies, property insurance, marine insurance, fire insurance, liability insurance and guarantee insurance. Personal insurance has changed in nature over the years. Insurance companies can offer a product which covers more than one risk of the insured. Currently, most personal insurance companies offer products that cover a combination of life, health and disability risks. Life benefits can be paid upon death, death within a set period of time and/or if the policy holder survives a set period of time as agreed in the contract. Health insurance contracts offer benefits to a policy holder when they fall sick.

It is therefore important for insurers to estimate the future lifetime of an insured life so that they may determine the amount of premium a life should pay by considering the number of payments to be made which in turn depends on the time of death of the insured life. This paper applies the Phase-type distribution for mortality to determine premiums payable to a whole life insurance policy.

1.2 Rationale

1.2.1 Multi-state Markov Models

At any particular time, a policy holder is in a given state and move from state to state with a given intensity. A multiple states Markov chain describes the process under which an individual moves from state to state in continuous time. The multiple state Markov Chain obeys the memoryless property of Markov processes. The chances that the process will move in certain state is not affected by previous states of the process but only on the current state.

Let us consider an insurance contract where a policyholder aged x taking a whole life insurance contract. The life can either be alive or dead. Consider the case where age is indexed according to calendar years.

In this study, the contract is entered when the policy holder is in state (X), i.e., alive and aged X and the contract ends when the policy holder enters the dead state (D). At any end of year, the life can either move to the next age, that is, state (X+1) or the dead state. If the policy is in state (X+1), that is, the life is aged (x+1) years, it can either move to state (X+2) or die. The same argument applies at ages. However, a life in the dead state cannot leave this state.

Traditionally for a whole life insurance policy, policyholders pay regular premiums for as long as the insured life is alive. For Markov process, the insured life is in state $X+j : j = 0,1,...,n$ where $x+n$ is the highest attainable age for individual in that particular population such that there are $n+1$ Markov states. A lump sum benefit is paid only when the policyholder enters the dead state. Therefore, for a life aged X entering a contract there is a probability of 1 that the Markov process starts in state (X) is 1 while the probability that it starts in any states $(X+j : j \geq 1)$ is 0.

1.2.2 Phase Type Distribution

Insurers pay benefit to a policyholder upon occurrence of the insured risk. For the multiple states Markov model explained in section 1.2 above, the sum assured will be paid when the insured life dies. Actuaries' work involve estimating claim time to be used together with the time value of money, in pricing insurance premiums.

The random variable underlining the phase-type distribution represents the time to absorption of a Markov Chain. An absorbing state a Markov process state such that if entered the process can not leave it again. That is, there is a zero probability of moving from the absorbing state to any other state, hence, the probability of stating in the absorbing state is 1 at any given time. Thus, absorption is the act of entering an absorbing state.

For a traditional whole life insurance policy the one absorbing state is death. However, there are two absorbing states for an endowment insurance policy, which are death and end of contract term. The probability of entering the end of contract term is zero during the contract period and is one at maturity.

The expected time of benefit payment can be estimated by determining the expected time to absorption of the multiple state Markov model and premium will therefore determined by discounting the sum assured with the estimated claim time. Premium can also be determined by discounting future payments Z times where Z is the random variable representing the future lifetime of the policy holder.

1.3 Definitions and Notations

1. **Insurance Policy**:These are legally binding contracts between two parties where one party (the insured) pays premium(s) to the other party (the insurer) who in turn promises to pay a benefit known as the **sum assured** in the case that the insured even occurs.

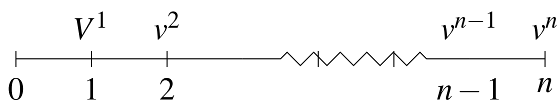
2. **Whole-life Insurance Contracts** are insurance policies where the benefits are payable at upon death of the insured life. Other Insurance contracts include term assurance, endowment insurance and pure endowment.

3. **Premiums** are amounts paid by the insured party to the insurer in order to cover for benefits and other insurance expenses. Premiums can be regular and level, lump sums, varying, deferred, among others.

4. **Present Value** is the current value of future payments while **expected present value** has adjustments to incorporate future risks and uncertainties. The major risk covered in a whole life insurance policy is death of the insured life.

5. **Present Value** is the current worth of a sum of money or stream of payments that will be made in the future. Most actuarial work involves determining present values than future values.

   consider the timeline below of a stream of n cash-flows made at times $t = 1, 2, ..., n$ and let i be applied interest rate per year, such that $v = \frac{1}{1+i} = 1 + i^{-1}$. Hence, the current value of each cash-flow made at time t is $C_t \times (1+i)^{-t} = C_t V^t$ where $C_t : t = 1, 2, ..3$ is the cash flow.



6. **Assurance** For a whole life insurance policy paying Ksh. 1 following the death of a life now aged x.

   <u>case a</u> For a case where benefit is paid at the end of the year that the insured dies. For instance, if the insured life dies at time $(n+i)$ where n is the number of full calendar years lived and i is a portion of the year such that $0 < i < 1$ and an interest rate of i p.a. is applied, then the number of years considered is $n+1$. The mean current value of such assurance benefit is:

$$A_x = E[V^{(n+1)}] \tag{1.3.0.1}$$

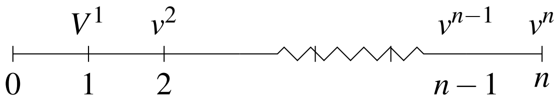
   <u>case b</u> consider a contract that pays the sum assured of Ksh. 1 immediately on death of a life (x) and let i be the rate of interest applied per year, such that the force of interest is σ p.a. and $V = e^{-\sigma}$. Let N denote the random variable representing the time death of life (x), the expected present value of such assurance is:

$$\begin{aligned} \bar{A}_x &= E[V^{(n)}] \\ &= E[e^{-\sigma \times n}] \end{aligned} \tag{1.3.0.2}$$

7. **Annuity** these are series of payments. Annuities can be level, certain, increasing, decreasing, deferred, made in advance or ordinary annuities whose first payment is made at the end of the first term. Consider level annuities paying Ksh.1 every year for n years,i.e., n payments will be made with the first payment being made in the first year.

   <u>Ordinary Annuity</u>

For this type of annuity the first payment is made at the end of the first year. The timeline below is a representation of such annuity where an interest of i p.a. is applied.

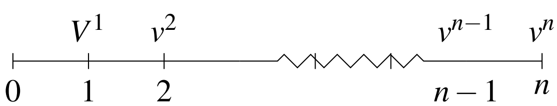


The present value of this annuity is:

$$\begin{aligned}
a_{\overline{n}|} &= v + v^2 + \ldots + v^{n-1} + v^n \\
&= v \times (1 + v^1 + \ldots + v^{n-2} + v^{n-1} \\
&= v\frac{1-v^n}{1-v} \\
&= \frac{1-v^n}{v^{-1}(1-v)} \\
&= \frac{1-v^n}{v^{-1}-1} \\
&= \frac{1-v^n}{1+i-1} \\
&= \frac{1-v^n}{i} \qquad \textbf{(1.3.0.3)}
\end{aligned}$$

<u>Annuity paid in Advance</u>

An annuity due or made in advance is such that the first payment is made at the beginning of the first year. The timeline below is a representation of such annuity where an interest of i p.a. is applied.



The present value of this annuity is:

$$\begin{aligned}
\ddot{a}_{\overline{n}|} &= 1 + v + v^2 + \ldots + v^{n-2} + v^{n-1} \\
&= \frac{1-v^n}{1-v} \\
&= \frac{1-v^n}{1-\frac{1}{1+i}} \\
&= \frac{1-v^n}{\frac{1+i-1}{1+i}} \\
&= \frac{1-v^n}{\frac{i}{1+i}} \\
&= \frac{1-v^n}{d} \qquad \textbf{(1.3.0.4)}
\end{aligned}$$

Where $d = \frac{i}{1+i}$

Annuity paid in continuously For an n years continuously payable annuity where an instantaneous interest rate of $\sigma p.a.$ is applied such that $\sigma = ln(1+i)$ where i is the annual effective rate of interest. The present value of such annuity can be derived through integration as follow:

$$\begin{aligned}
\bar{a}_{\overline{n}|} &= \int_{t=0}^{n} v^t dt \\
&= \left[\frac{v^t}{lnv}\right]_0^n \\
&= \frac{v^n - v^0}{lnv} \\
&= \frac{v^n - 1}{ln\left(\frac{1}{1+i}\right)} \\
&= \frac{v^n - 1}{ln(1) - ln(1+i)} \\
&= \frac{v^n - 1}{0 - ln(1+i)} \\
&= \frac{1 - v^n}{ln(1+i)} \\
&= \frac{1 - v^n}{\sigma}
\end{aligned} \quad \textbf{(1.3.0.5)}$$

Evaluating Life Annuities

Ordinary Annuity

Consider an annuity paying Ksh.1 annually for the whole of life of a person now aged x.If payments are made at the end of each year, then, the number of payments done equals the number of whole years lived by (x) , in that, if the person lives $n+i$ where n is the number of whole years lived and i is a portion of a year such that $0 < i < 1$ then n payments will be done. This implies that the life lives to age $x+n$ but dies before age $x+n+1$.

Let N be the random variable representing the future lifetime of a person now aged (x). The expected present value of the life annuity is

$$\begin{aligned}
a_{x:\overline{n}|} &= E[a_{\overline{n}|}] \\
&= E[\frac{1 - v^n}{i}] \\
&= \frac{1 - E[v^n]}{i} \\
&= \frac{1 - A_x}{i}
\end{aligned} \quad \textbf{(1.3.0.6)}$$

Annuity Due

Consider an annuity paying Ksh.1 at the start of each year for the whole of life of a person now aged x. Premiums usually take a form of this annuity. The number of payments done equals the number of whole years lived by (x) plus 1 , in that, if the person lives $n+i$ where n is the number of whole years lived and i is a portion of a year such that $0<i<1$ then $n+1$ payments will be done.

Let N denote the random variable representing the future lifetime of a person now aged (x). The expected present value of this life annuity is

$$\begin{aligned}\ddot{a}_{x:\overline{n}|} &= E[\ddot{a}_{\overline{n}|}] \\ &= E[\frac{1-v^n}{d}] \\ &= \frac{1-E[v^n]}{d} \\ &= \frac{1-A_x}{d}\end{aligned} \qquad \textbf{(1.3.0.7)}$$

Annuity Paid Continuously

Consider an annuity paying Ksh.1 p.a as long as a person now aged x lives. Let interest rate be i per year, translating to an instantaneous interest of σ such that $i=1-e^{\sigma}$, is applied. Let N be the random variable representing the future lifetime of a life now aged x. The expected present value of this annuity is given as:

$$\begin{aligned}\bar{a}_{x:\overline{n}|} &= E[\bar{a}_{\overline{n}|}] \\ &= E[\frac{1-v^n}{\sigma}] \\ &= \frac{1-E[v^n]}{\sigma} \\ &= \frac{1-A_x}{\sigma}\end{aligned} \qquad \textbf{(1.3.0.8)}$$

8. **Markov Chain** A Markov chain is a stochastic process such as the probability of an event occuring depends only on the current state of the process. This property is known as the memoryless markov property. Consider a markov chain $\{X_n, n = 0,1,2,..\}$, the Markov property is represented as:

$$Prob\{X_{n+1}=x_{n+1}|X_1=x_1,X_2=x_2,...X_n=x_n\}=Prob\{X_{n_+1}=x_{n+1}|X_n=x_n\}$$

9. **Absorbing State** this is a Markov state such that if entered the process cannot leave it. Examples of absorbing states in actuarial science include ruin, death, permanent disability and retirement (in Pension) among others. Those states that are not absorbing states are **transient states** in that the process can enter and leave the state.

10. **Transition matrix** this is an array representing the rate at which a process leaves from state to state. The rate of stay in each state is also represented in such a matrix. A stochastic matrix contains the transition rates that a Markov process has when it

moves between states or stays in each state in a discrete time. Let p_{ij} be the transition rate from state i to state j and that the process has N states. The following two properties are true about a transition matrix and transition rates:
i. Rows of a transition matrix add up to 1

$$\sum_i p_{ij} = 1$$
$$p_{ii} = 1 - \sum_{i \neq j} p_{ij} \tag{1.3.0.9}$$

ii. Transition rates are probabilities, hence, they take values between 0 and 1

$$0 < p_{ij} < 1 \tag{1.3.0.10}$$

11. **A Generator Matrix**, also known as the infinitesimal generator matrix, is an array representing the transition rates of a continuous time Markov chain as it moves from state to state. Let $q_i j$ be the rate at which the process leaves state i to state j within a very short period of time. This is often called intensity rate. The following two properties of intensity rates apply for generator matrices:
i. Rows of a generator matrix add up to 0

$$\sum_i q_{ij} = 0$$
$$q_{ii} = -\sum_{i \neq j} q_{ij} \tag{1.3.0.11}$$

ii. intensity rates between states are non-negative

$$q_{ij} \geq 0 \tag{1.3.0.12}$$

1.4 Problem Statement

Recently there has been increasing interest in mortality modeling and projection due to the mortality improvement in the recent years and the consequent adverse nancial impact on pension plans and annuity business. For instance, the world life expectancy of male, females and both sexes has witnessed an increase of 3.14%, 3.36%, and 3.24% respective (Organization, 2015)(Organization, 2020). This increasing trend is expected to continue, hence, the need to have a model that does not have a maximum attainable age. Past attempts on mortality projection have often underestimated the overall mortality improvement. The Lee-carter model suggests that there is no limiting age, however, this model is not sufficient for computation of actuarial functions (Lee & Carter, 1992). On the other hand, the De Moivre model which was purposely built for computation of annuity functions, and possibly other actuarial functions, requires a limiting age (de Moivre, 1725).

Similarly, the life tables functions are also computed for up to the limiting age (Graunt, 1973). A life table is built such that there is a certain age beyond which individuals do not live. The probability of survival beyond this particular age is assumed to be zero. For instance, the maximum attainable age for the AM92 tables used for examination by the Institute and Faculty of actuaries is 120.

There is therefore need to develop a mortality model that fits mortality data. The same model should takes into account the biological aging process and can flexibly be adjusted to incorporate medical opinion. The model should also be suitable computation of actuarial functions. This paper proposes the use of phase-type distribution to describe a physiological aging process of a human body.

A phase type distribution is applicable for approximation nearly any other distribution positive distribution (Asmussen, 1989). The random variable underlying this distribution takes greater than 0 on the real line for a continuous type and takes indexing values starting from 0, $i = 0, 1, 2, \ldots$, for a discrete type, hence, satisfying the requirement of no-limiting age. Also, expected values of functions can be determined for phase-type distributions, thus, the requirement of computing actuarial functions is satisfied. Besides, the phase type distribution has a tractable property and have explicit solutions of exponential functions.

1.5 Study Objectives

1.5.1 Overall Study Objective

To apply the Phase-type Distribution in determining Premiums payable for a while life insurance policy.

1.5.2 Specific Objectives

1. To examine the applicability of phase type distribution in computation of actuarial functions.

2. To develop a distribution that fits mortality data and physiological aging process .

1.6 Significance of Study

This is an extensive study of Phase-Type distribution and its application to actuarial science. Explanations of derivation of moments including the mean and variances can be used for statistics lectures at tertiary levels. Actuarial students as well as insurance companies can make use of this paper for premiums calculations.

Literature of past mortality models and past applications of phase type distribution will be reviewed in Chapter 2 of this project. Chapter 3 introduces a Markov Chain with one absorbing state and later described the Discrete Phase-Type distribution in details. The continuous Time Markov chain and Continuous Time Phase-Type distribution are discussed in chapter 4. The paper then applies the continuous phase type distribution to AM92 life table functions to estimate whole life actuarial functions including mean present values of assurances, annuities and net premiums. The paper ends in Chapter 6 by discussing the results, suggesting future studies and making a conclusion .

2 Literature Review

Several models have been built in a bid to explain mortality phenomenon. It is accepted that the rate of death increases as life ages. In 1725, Abraham De Moivre developed a mortality model known as *"De Moivre Law*. He built this model to be used in annuity and pension cost calculations. The model has since been used in most actuarial calculations. In his 1731 book, he suggested that the probability of death increases with age while survivorship decreases (de Moivre, 1725) such that $S(x) = 1 - \frac{x}{\omega}$ where $s(x)$ is the probability that a new born will be alive to least to age x, while ω is the highest attainable age, that is, the limiting age. From this model, the probability, ${}_tp_x$, that a life now aged x will survive at least t years is determined as ${}_tp_x = \frac{S(x+t)}{S(x)} = \frac{1-\frac{x+t}{\omega}}{1-\frac{x}{\omega}} = \frac{\omega-(x+t)}{\omega-x}$. The probability that a person now aged x dies before age $x+t$ can also be determined from the *DeMoivre Law* as ${}_tq_x = \frac{S(x)-S(x+t)}{S(x)} = \frac{(1-\frac{x}{\omega})-(1-\frac{x+1}{\omega})}{1-\frac{x}{\omega}} = \frac{t}{w-x}$. Hence, the probability that a life aged ω survives beyond this age is zero.

Benjamin Gompertz in his 1825 actuarial work proposed the *Gompertz Law* which suggests that the instantaneous rate of death, also known as hazard, is dependant on age and that (1) hazard increases as age increases, however (2) the rate of change in the hazard is slower at older ages (Gompertz, 1825). The hazard rate $\mu(x,a,b)$ is an exponential function of hazard at birth a, increase in hazard b and current age x of a life in that $\mu(x,a,b) = ae^{bx}$.

Later in 1860, William Makeham proposed an additional age independent parameter to the Gompertz model, which resulted in the *Gompertz-Makeham model* (Makeham, 1860). The resulting hazard function is the sum of the Gompertz component ae^{bx} and the Makeham componet λ such that $\mu(x,a,b,\lambda) = ae^{bx} + \lambda$.

Later on (Makeham, 1860) proposed an eight component mortality model, $\mu_x = A^{(x+B)^C} + De^{E(logxlogF)^2} + GH^x$ where there are three parameters A, B and C to describe child mortality, three others D, E and F to describe a very flexible accident hump that typically occurs in young adulthood, and finally two parameters G and H to describe mortality at older ages. The main disadvantage of this model is that in its traditional form is difficult to fit and it does not account for uncertainty.

In 1992 Ronald Lee and Lawrence Carter built a simple model for describing the change in total mortality as a function of a single time parameter, k_t (Lee & Carter, 1992).The model is structured as $\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$ where $m_{x,t}$ is the central rate of mortality of a life aged x in year t, a_x is a mortality parameter that depends on age only, b_x is a parameter describing the deviation in mortality for lives aged x as the time component k_t varies and $\varepsilon_{x,t}$ is the error term.

The *Lee-Carter* has been widely used for mortality projections. The main result of the model is that human mortality is improving in that it will reach a point where hazard will be negligible. This is partly attributed to improvement in the health sector. Also, the longevity risk will be most evident in older population since their life expectancy is higher than that of younger population.

3 Discrete Phase-Type Distribution

3.1 A Markov Chain with one Absorbing State

3.1.1 Preliminaries

In this section, we will discuss the fundamental characteristics of Markov chains with one absorbing state. We shall define the transition matrix and highlight the fundamental properties of Markov Chains with one absorbing state.

3.1.2 A Markov Chain with One Absorbing State

Let

$$S = \{0, 1, 2, \ldots\}$$

be a state space with 0, being an absorbing state and the other m states are transient. The transition matrix can be partitioned into a 2x2 block as follows:

$$P = \left[\begin{array}{c|ccccc} 1 & 0 & 0 & 0 & \cdots & 0 \\ \hline P_{1,0} & P_{1,1} & P_{1,2} & P_{1,3} & \cdots & P_{1,m} \\ P_{2,0} & P_{2,1} & P_{2,2} & P_{2,3} & \cdots & P_{2,m} \\ P_{3,0} & P_{3,1} & P_{3,2} & P_{3,3} & \cdots & P_{3,m} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{m,0} & P_{m,1} & P_{m,2} & P_{m,3} & \cdots & P_{m,m} \end{array}\right]$$

$$P = \begin{bmatrix} 1 & \underline{O} \\ \underline{U} & T \end{bmatrix} \tag{3.1.2.1}$$

Where;
P is $(m \times m) \times (m \times 1)$ matrix 1 is 1×1
O' is 1×1 m zero row vector
U is $m \times 1$ transition matrix from a transient state to a transient state.

Let

$$\underline{e} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{3.1.2.2}$$

which is an m x 1 column vector consisting of 1.

$$\therefore T\underline{e} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \vdots & & & \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11}+p_{12}+\dots+p_{1m} \\ p_{21}+p_{22}+\dots+p_{2m} \\ \vdots \\ p_{m1}+p_{m2}+\dots+p_{mm} \end{bmatrix}$$

Since each row of the transition matrix adds up to 1, then;

$$\begin{bmatrix} p_{10} \\ p_{20} \\ \vdots \\ p_{m0} \end{bmatrix} + \begin{bmatrix} p_{11}+p_{12}+\dots+p_{1m} \\ p_{21}+p_{22}+\dots+p_{2m} \\ \vdots \\ p_{m1}+p_{m2}+\dots+p_{mm} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

i.e.;

$$\underline{U} + T\underline{e} = \underline{e}$$

$$\therefore \underline{U} = \underline{e} - T\underline{e}$$

$$\underline{U} = [I - T]\underline{e}$$

where I is an identity matrix.

$$\therefore P = \begin{bmatrix} 1 & \underline{O'} \\ e - T\underline{e} & T \end{bmatrix}$$

The initial probability distribution vector is given by;

$$\underline{\beta'} = \{\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_m\} = \{\alpha_0, \underline{\alpha_0'}\}$$

where;

$$\underline{\alpha'} = \alpha_1, \alpha_2, \dots, \alpha_m$$

and

$$\alpha_0 + \alpha_1 + \alpha_2 + \dots + \alpha_m = 1 \quad \textbf{(3.1.2.3)}$$

$$P^2 = \begin{bmatrix} 1 & \underline{O}' \\ \underline{e} - T\underline{e} & T \end{bmatrix} \begin{bmatrix} 1 & \underline{O}' \\ \underline{e} - T\underline{e} & T \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 1 & \underline{O}' \\ B_2 & T^2 \end{bmatrix}$$

Where

$$B_2 = (\underline{e} - T\underline{e}) + T(\underline{e} - T\underline{e})$$

$$= \underline{e} - T\underline{e} + T\underline{e} - T^2\underline{e}$$

$$= \underline{e} - T^2\underline{e}$$

$$\therefore P^2 = \begin{bmatrix} 1 & \underline{O}' \\ \underline{e} - T^2\underline{e} & T^2 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 1 & \underline{O}' \\ \underline{e} - T^2\underline{e} & T \end{bmatrix} \begin{bmatrix} 1 & \underline{O}' \\ \underline{e} - T\underline{e} & T \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 1 & \underline{O}' \\ B_3 & T^3 \end{bmatrix}$$

Where

$$B_3 = \underline{e} - T^2\underline{e} + T^2\underline{e} - T^3\underline{e}$$

$$= \underline{e} - T^3\underline{e}$$

$$\therefore P^3 = \begin{bmatrix} 1 & \underline{O}' \\ \underline{e} - T^3\underline{e} & T^3 \end{bmatrix}$$

By induction;

$$P^n = \begin{bmatrix} 1 & \underline{O}' \\ \underline{e} - T^n\underline{e} & T^n \end{bmatrix} \quad \textbf{(3.1.2.4)}$$

Alternatively, using

$$P = \begin{bmatrix} 1 & \underline{o}' \\ \underline{U} & T \end{bmatrix}$$

Then;

$$P^2 = \begin{bmatrix} 1 & \underline{o}' \\ \underline{U} & T \end{bmatrix} \begin{bmatrix} 1 & \underline{o}' \\ \underline{U} & T \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \underline{o}' \\ B_2 & T^2 \end{bmatrix}$$

Where;

$$B_2 = \underline{U} + T\underline{U}$$

$$P^3 = \begin{bmatrix} 1 & \underline{o}' \\ B_2 & T^2 \end{bmatrix} \begin{bmatrix} 1 & \underline{O}' \\ \underline{U} & T \end{bmatrix} = \begin{bmatrix} 1 & \underline{O}' \\ B_3 & T^3 \end{bmatrix}$$

Where;

$$B_3 = B_2 + T^2\underline{U}$$

$$= \underline{U} + T\underline{U} + T^2\underline{U}$$

$$P^4 = \begin{bmatrix} 1 & \underline{o}' \\ B_3 & T^3 \end{bmatrix} \begin{bmatrix} 1 & \underline{O}' \\ \underline{U} & T \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \underline{O}' \\ B_4 & T^4 \end{bmatrix}$$

Where;

$$B_4 = B_3 + T^3\underline{U}$$

$$= \underline{U} + T\underline{U} + T^2\underline{U} + T^3\underline{U}$$

$$= (I + T + T^2 + T^3)\underline{U}$$

$$= (T^0 + T^1 + T^2 + T^3)\underline{U}$$

$$= (\sum_{k=0}^{3} T^k)\underline{U} \tag{3.1.2.5}$$

By induction

$$P^n = \begin{bmatrix} 1 & \underline{O'} \\ B_n & T^n \end{bmatrix}$$

Where

$$B_n = \left(\sum_{k=0}^{n-1} T^k\right)\underline{U} \tag{3.1.2.6}$$

consider;

$$\sum_{k=o}^{n-1} T^k = T^0 + T^1 + T^1 + \ldots + T^{n-1}$$
$$= I + T + T^2 + \ldots + T^n$$
$$\therefore (I-T)(I+T+T^2+\ldots+T^{n-1}) = I - T^n$$
$$\therefore (I-T)^{-1}(I-T)(I+T+T^2+\ldots+T^{n-1}) = (I-T)^{-1}(I-T^n)$$

i.e

$$T + T^2 + \ldots + T^{n-1} = (I-T)^{-1}(I-T^n)$$
$$\therefore \sum_{k=0}^{n-1} T^k = (I-T)^{-1}(I-T^n)$$
$$\therefore B_n = (I-T)^{-1}(I-T^n)\underline{U}$$
$$\therefore P^n = \begin{bmatrix} 1 & \underline{O'} \\ (I-T)^{-1}(I-T^n)U & T^n \end{bmatrix}$$
$$\therefore \lim_{n\to\infty} P^n = \lim_{n\to\infty} \begin{bmatrix} 1 & \underline{O'} \\ (I-T)^{-1}(I-T^n)\underline{U} & T^n \end{bmatrix}$$
$$= \lim_{n\to\infty} \begin{bmatrix} 1 & \underline{O'} \\ (I-T)^{-1}(I-\lim_{n\to\infty} T^n)\underline{U} & \lim_{n\to\infty} T^n \end{bmatrix} \tag{3.1.2.7}$$

From;

$(I-T)(I+T+T^2+\ldots+T^{n-1}) = I-T^n$

take the limit, i.e.;

$\lim_{n\to\infty}(I-T)(I+T+T^2+\ldots+T^{n-1}) = I-\lim_{n\to\infty}T^n$

$\therefore (I-T)(I+T+T^2+\ldots+\lim_{n\to\infty}T^{n-1}) = I-\lim_{n\to\infty}T^n$

i.e.;

$(I-T)(I+T+T^2+\ldots) = I-\lim_{n\to\infty}T^n$

$\therefore I = I-\lim_{n\to\infty}T^n$

$\lim_{n\to\infty}T^n = 0$

$$\therefore \lim_{n\to\infty}P^n = \begin{bmatrix} 1 & \underline{O'} \\ & (I-T)^{-1}(I-\underline{O'})\underline{U} \quad \underline{O} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \underline{O'} \\ (I-T)^{-1}\underline{U} & \underline{O} \end{bmatrix} \tag{3.1.2.8}$$

Since

$\underline{U} = (I-T)\underline{e}$

then

$(I-T)^{-1}\underline{U} \quad = (I-T)^{-1}(I-T)\underline{e} = \underline{e}$

$$\therefore \lim_{n\to\infty}P^n = \begin{bmatrix} 1 & \underline{O'} \\ \underline{e} & \underline{O} \end{bmatrix} \tag{3.1.2.9}$$

3.1.3 Interpretation of $(I-T)^{-1}$

$(I-T)^{-1}$ is called the fundamental matrix which is the average number of times a system is in transient state E_j from a transient state E_i.

3.2 Discrete Phase-Type (PH) Distribution

3.2.1 Preliminaries

In the previous chapter we discussed Markov chains with one absorbing state. In this chapter, we are going to introduce the idea of discrete time Phase Type distribution and apply it to Markov Chain. We shall also discuss fundamental properties of the Discrete Phase-Type Distribution.
Recall: Any Markov chain consists of transaction probability or matrix and the initial distribution.

3.2.2 Discrete Phase-Type Distribution

A fundamental characteristic of an absorbing Markov Chain is the time to absorption, i.e.; the number of steps the process takes to reach an absorbing state. The distribution of the time to absorption is called Phase-Type (PH) distribution. Each transient state is called a phase.
Let

$$f(z) = Prob\{Z = z\} \tag{3.2.2.1}$$

Where Z is the random variable denoting the number of steps the process takes to reach the absorbing state.

$\therefore f(z) =$ the probability of being absorbed in z steps

$=$ the probability of starting at any transient state, then moving over transient states in $z-1$ steps and finally getting absorbed.

$= \underline{\alpha}' T^{z-1} \underline{U}$, where $z = 1, 2, 3, \ldots$ and $f(0) = \alpha_0$

(3.2.2.2)

$F(z) = Prob\{Z \leq z\}$

=the probability that absorption occurs at or before time z.

$\therefore 1 - F(z) = Prob\{Z > z\}$

$=$ the probability that absorption does not occur before time z having started at a transient state.

$= \underline{\alpha}' T^z \underline{e}$

$\therefore F(z) = 1 - \underline{\alpha}' T^z \underline{e}$, for $z = 0, 1, 2, \ldots$

(3.2.2.3)

Remark 3.2.1. $f(0) = \alpha_0$

3.2.3 Probability Generating Function of a PH Distribution

Theorem 3.2.2. *Let Z be a discrete phase type random variable, T is the underlining $m \times m$ transition matrix of transient state and α_0 and $\underline{\alpha}'$ are the initial probabilities of the absorbing state and transient states respectively. The Probability Generating Function of Z is then given by;*

$$\begin{aligned} G(s) &= \sum_{z=0}^{\infty} f(z)s^z \\ &= \alpha_0 + \underline{\alpha}' s(I - sT)^{-1}\underline{U} \end{aligned} \tag{3.2.3.1}$$

Proof.

$$\begin{aligned} G(s) &= \sum_{z=0}^{\infty} f(z)s^z \\ &= f(0) + \sum_{z=1}^{\infty} f(z)s^z \\ &= \alpha_0 + \sum_{z=1}^{\infty} \underline{\alpha}' T^{z-1}\underline{U}s^z \\ &= \alpha_0 + \underline{\alpha}'(\sum_{z=1}^{\infty} T^{z-1})\underline{U}s^z \\ &= \alpha_0 + \underline{\alpha}' s(I + sT + (sT)^2 + \ldots)\underline{U} \\ &= \alpha_0 + \underline{\alpha}' s(I - sT)^{-1}\underline{U} \\ \therefore G(s) &= \alpha_0 + \underline{\alpha}' s(I - sT)^{-1}\underline{U} \end{aligned} \tag{3.2.3.2}$$

□

Before differentiating G(s) we need the following;

Lemma 3.2.3. *Let A be a non-singular square matrix whose elements are functions of a scalar S. Then;*

$$\frac{d}{ds}A^{-1} = -A^{-1}\frac{dA}{ds}A^{-1}$$

Proof. From

$$A^{-1}A = I$$

$$\frac{d}{ds}A^{-1}A = \frac{d}{ds}I$$

i.e.;

$$\left(\frac{dA^{-1}}{ds}\right)A + A^{-1}\frac{dA}{ds} = 0$$

$$\therefore \left(\frac{dA^{-1}}{ds}\right)A = -A^{-1}\frac{dA}{ds}$$

$$\therefore \frac{d}{ds}A^{-1} = -A^{-1}\frac{dA}{ds}A^{-1} \tag{3.2.3.3}$$

□

Lemma 3.2.4. *Let I be an identity matrix, T be a square matrix, and s be a scalar;*

1. $(I-sT)^{-1}T = T(I-sT)^{-1}$
   and in general

2. $(I-sT)^{-k}T = T(I-sT)^{-k}$ *where* $K = 1,2,3,\dots$
   Thus $(I-sT)^{-k}$ *and* T *are commutative.*

Proof.

$$
\begin{aligned}
(I-sT)^{-1}T &= (I-sT)^{-1}\frac{sT}{s} \\
&= \frac{1}{s}(I-sT)^{-1}sT \\
&= \frac{1}{s}(I-sT)^{-1}\{I-I+sT\} \\
&= \frac{1}{s}(I-sT)^{-1}\{I-(I-sT)\} \\
&= \frac{1}{s}\{(I-sT)^{-1}-I\} \\
&= \frac{1}{s}\{(I-sT)^{-1}-(I-sT)(I-sT)^{-1}\} \\
&= \frac{1}{s}\{I-(I-sT)\}(I-sT)^{-1} \\
&= \frac{1}{s}.sT(I-sT)^{-1} \\
&= T(I-sT)^{-1}
\end{aligned}
\tag{3.2.3.4}
$$

In general;

$$
\begin{aligned}
(I-sT)^{-k}T &= (I-sT)^{-1}\frac{sT}{s} \\
&= \frac{1}{s}(I-sT)^{-k}\{I-(I-sT)\} \\
&= \frac{1}{s}\{(I-sT)^{-k}-(I-sT)^{-k+1}\} \\
&= \frac{1}{s}\{(I-sT)^{-k+1-1}-(I-sT)^{-k+1}\} \\
&= \frac{1}{s}\{(I-sT)^{-1}-I\}(I-sT)^{-k+1} \\
&= \frac{1}{s}\{(I-sT)^{-1}-(I-sT)(I-sT)^{-1}\}(I-sT)^{-k+1}
\end{aligned}
\tag{3.2.3.5}
$$

□

Theorem 3.2.5. *For $PH(\underline{\alpha},T)$ the first, second, third and k^{th} moment generating functions are;*

1. $G'(s)=\underline{\alpha'}\{(I-sT)^{-1}+s(I-sT)^{-2}T\}\underline{U}$
2. $G''(s)=2\underline{\alpha'}\{(I-sT)^{-2}T+s(I-sT)^{-3}T^2\}\underline{U}$
3. $G'''(s)=3!\underline{\alpha'}\{(I-sT)^{-3}T^2+s(I-sT)^{-4}T^3\}\underline{U}$
4. $G^k(s)=k!\underline{\alpha'}\{(I-sT)^{-k}T^{k-1}+s(I-sT)^{-(k+1}T^k\}\underline{U}$

Proof. .

1.

$$\begin{aligned}G'(s)&=\frac{d}{ds}\left\{\alpha_0+\underline{\alpha'}s\,(I-sT)^{-1}\,\underline{U}\right\}\\&=\underline{\alpha'}\left\{\frac{d}{ds}s\,(I-sT)^{-1}\right\}\underline{U}\\&=\underline{\alpha'}\left\{(I-sT)^{-1}+s\frac{d}{ds}\,(I-sT)^{-1}\right\}\underline{U}\end{aligned}$$

Applying Lemma 1,

$$\begin{aligned}G'(s)&=\underline{\alpha'}\left\{(I-sT)^{-1}-s(I-sT)^{-1}\left[\frac{d}{ds}(I-sT)\right](I-sT)^{-1}\right\}\underline{U}\\&=\underline{\alpha'}\left\{(I-sT)^{-1}+s(I-sT)^{-1}T(I-sT)^{-1}\right\}\underline{U}\end{aligned}$$

Applying Lemma 2,

$$G's=\underline{\alpha'}\left\{(I-sT)^{-1}+s(I-sT)^{-2}T\right\} \qquad \textbf{(3.2.3.6)}$$

2.

$$\begin{aligned}G''(s)&=\underline{\alpha'}\left\{\frac{d}{ds}(I-sT)^{-1}+\frac{d}{ds}s(I-sT)^{-2}T\right\}\underline{U}\\&=\underline{\alpha'}\left\{-(I-sT)^{-1}\frac{d}{ds}s(I-sT)^{-1}+(I-sT)^{-2}T+s\frac{d}{ds}\left[(I-sT)^{-1}\right]^2T\right\}\underline{U}\\&=\underline{\alpha'}\left\{(I-sT)^{-1}T(I-sT)^{-1}+(I-sT)^{-2}T+2s(I-sT)^{-1}\frac{d}{ds}(I-sT)^{-1}T\right\}\underline{U}\\&=\underline{\alpha'}\left\{(I-sT)^{-2}T+(I-sT)^{-2}T+2s(I-sT)^{-1}\left[-(I-sT)^{-1}\frac{d}{ds}(I-sT).(I-sT)^{-1}\right]T\right\}\underline{U}\\&=\underline{\alpha}\left\{2(I-sT)^{-2}T+2s(I-sT)^{-2}T(I-sT)^{-1}T\right\}\underline{U}\\&=\underline{\alpha}\left\{2(I-sT)^{-2}+2s(I-sT)^{-3}T^2\right\}\underline{U}\\&=2\underline{\alpha}\{(I-sT)^{-2}T+s(I-sT)^{-3}T^2\}\underline{U}\end{aligned}$$

3.

$$
\begin{aligned}
G'''(s) &= 2\underline{\alpha'}\left\{\frac{d}{ds}[(I-sT)^{-1}]^2T + \frac{d}{ds}s[(I-sT)^{-1}]^3T^2\right\}\underline{U} \\
&= 2\underline{\alpha'}\left\{-2(I-sT)^{-1}\frac{d}{ds}(I-sT)(I-sT)^{-1}T + (I-sT)^{-3}T^2 + s\frac{d}{ds}s\left[(I-sT)^{-1}\right]^3T^2\right\}\underline{U} \\
&= 2\underline{\alpha'}\left\{-2(I-sT)^{-1}(I-sT)^{-1}\left[\frac{d}{ds}(I-sT)^{-1}\right](I-sT)^{-1}T + (I-sT)^{-3}T^2\right\}\underline{U} \\
&+ 2\underline{\alpha'}\left\{3s(I-sT)^{-2}\left[\frac{d}{ds}(I-sT)^{-1}\right]T^2\right\}\underline{U} \\
&= 2\underline{\alpha'}\left\{2(I-sT)^{-2}T(I-sT)^{-1}T + (I-sT)^{-3}T^2\right\} \\
&+ 2\underline{\alpha'}\left\{3s(I-sT)^{-2}(I-sT)^{-1}T.(I-sT)^{-1}T.(I-sT)^{-1}T^2\right\} \\
&= 2\underline{\alpha'}\left\{2(I-sT)^{-3}T^2 + (I-sT)^{-3}T^2 + 3s(I-sT)^{-3}T(I-sT)^{-1}T^2\right\}\underline{U} \\
&= 2\underline{\alpha'}\left\{3(I-sT)^{-3}T^2 + 3s(I-sT)^{-4}T^3\right\}\underline{U} \\
\therefore G'''(s) &= 3!\underline{\alpha'}\left\{(I-sT)^{-3}T^2 + 3s(I-sT)^{-4}T^3\right\}\underline{U}
\end{aligned}
\tag{3.2.3.7}
$$

4. By induction; assume

$$G^{(k-1)}(s) = (k-1)!\underline{\alpha'}\{(I-sT)^{-(k-1)}T^{k-2} + s(I-sT)^{-k}T^{k-1}\}\underline{U}$$

to be true. Then;

$$
\begin{aligned}
G^{(k)}(s) &= (k-1)!\underline{\alpha'}\{\frac{d}{ds}(I-sT)^{-(k-1)}T^{k-2} + \frac{d}{ds}s(I-sT)^{-k}T^{k-1}\}\underline{U} \\
&= (k-1)!\underline{\alpha'}\{\frac{d}{ds}[(I-sT)^{-1}]^{k-1}T^{k-2} + (I-sT)^{-k}T^{k-1} + s\frac{d}{ds}[(I-sT)^{-1}]^kT^{k-1}\}\underline{U} \\
&= (k-1)!\underline{\alpha'}\{(k-1)(I-sT)^{-k+2}[\frac{d}{ds}(I-sT)^{-1}]T^{k-2} + (I-sT)^{-k}T^{k-1} \\
&+ sk(I-sT)^{-(k-1)}[\frac{d}{ds}(I-sT)^{-1}]T^{k-1}\}\underline{U} \\
\therefore G^{(k)}(s) &= (k-1)!\underline{\alpha'}\{(k-1)(I-sT)^{-k+2}(I-sT)^{-1}T(I-sT)^{-1}T^{k-2} + (I-sT)^{-k}T^{k-1} \\
&+ sk(I-sT)^{-(k+1)}(I-sT)^{-1}T(I-sT)^{-1}T^{k-1}\}\underline{U} \\
&= (k-1)!\underline{\alpha'}\{(k-1)(I-sT)^{-k+2-1-1}T^{k-2+1} + (I-sT)^{-k}T^{k-1} \\
&+ sk(I-sT)^{-k+1-1-1}T^{k-1+1}\}\underline{U} \\
&= (k-1)!\underline{\alpha'}\{k(I-sT)^{-k}T^{k-1} + sk(I-sT)^{-(k+1)}T^k\}\underline{U} \\
&= k!\underline{\alpha'}\{(I-sT)^{-k}T^{k-1} + s(I-sT)^{-(k+1)}T^k\}\underline{U}
\end{aligned}
\tag{3.2.3.8}
$$

□

Corollary 3.2.6.

$$\begin{aligned}
G'(1) &= E(z) = \underline{\alpha}'\{(I-sT)^{-1}+(I-sT)^{-2}\}\underline{U}\\
G''(1) &= E[z(z-1)] = 2\underline{\alpha}'\{(I-sT)^{-2}T+(I-sT)^{-3}T^2\}\underline{U}\\
G'''(1) &= E[z(z-1)(z-2)] = 3!\underline{\alpha}'\{(I-sT)^{-3}T^2+(I-sT)^{-4}T^3\}\underline{U}\\
G^{(k)} &= E[z(z-1)(z-2)\ldots(z-(k+1)] = k!\underline{\alpha}'\{(I-sT)^{-k}T^{k-1}+(I-sT)^{-(k+1)}T^k\}\underline{U}
\end{aligned}$$

$$\begin{aligned}
E(z) = G'(1) &= \underline{\alpha}\{(I-T)^{-1}+(I-T)^{-2}T\}\underline{U}\\
&= \underline{\alpha}(I-T)^{-1}\{I+(I-T)^{-1}T\}\underline{U}\\
&= \underline{\alpha}(I-T)^{-1}\{I+T(I-T)^{-1}\}\underline{U}\\
&= \underline{\alpha}(I-T)^{-1}\{(I-T)(I-T)^{-1}+T(I-T)^{-1}\}\underline{U}\\
&= \underline{\alpha}(I-T)^{-1}\{(I-T)+T\}(I-T)^{-1}\underline{U}\\
&= \underline{\alpha}(I-T)^{-1}\underline{e}
\end{aligned} \tag{3.2.3.9}$$

$$\begin{aligned}
E[z(z-1)] = G''(1) &= 2\underline{\alpha}'\{(I-T)^{-2}T+(I-T)^{-3}T^2\}\underline{U}\\
&= 2\underline{\alpha}'(I-T)^{-2}T\{I+(I-T)^{-1}T\}\underline{U}\\
&= 2\underline{\alpha}'(I-T)^{-2}T\{I+T(I-T)^{-1}\}\underline{U}\\
&= 2\underline{\alpha}'(I-T)^{-2}T\{(I-T)(I-T)^{-1}+T(I-T)^{-1}\}\underline{U}\\
&= 2\underline{\alpha}(I-T)^{-2}\{(I-T)+T\}(I-T)^{-1}\underline{U}\\
&= 2\underline{\alpha}(I-T)^{-2}T\underline{e}
\end{aligned} \tag{3.2.3.10}$$

$$\begin{aligned}
E[z(z-1)(z-2)] = G'''(1) &= 3!\underline{\alpha}'(I-T)^{-3}\{T^2+(I-T)^{-1}T^3\}\underline{U}\\
&= 3!\underline{\alpha}'(I-T)^{-3}\{T^2+T^3(I-T)^{-1}\}\underline{U}\\
&= 3!\underline{\alpha}'(I-T)^{-3}T^2\{I+T(I-T)^{-1}\}\underline{U}\\
&= 3!\underline{\alpha}'(I-T)^{-3}T^2\{(I-T)(I-T)^{-1}+T(I-T)^{-1}\}\underline{U}\\
&= 3!\underline{\alpha}(I-T)^{-3}T^2\{(I-T)+T\}(I-T)^{-1}\underline{U}\\
&= 3!\underline{\alpha}(I-T)^{-3}T^2\underline{e}
\end{aligned} \tag{3.2.3.11}$$

In general;

$$
\begin{aligned}
E[z(z-1)(z-2)\dots(z-k+1)] = G^{(k)}(1) &= k!\underline{\alpha'}\{(I-T)^{-k}T^{k-1} + (I-T)^{-(k+1)}T^{K}\}\underline{U} \\
&= k!\underline{\alpha'}(I-T)^{-k}\{T^{(k-1)} + (I-T)^{-1}T^{k}\}\underline{U} \\
&= k!\underline{\alpha'}(I-T)^{-k}\{T^{k-1} + T^{k}(I-T)^{-1}\}\underline{U} \\
&= k!\underline{\alpha'}(I-T)^{-k}T^{k-1}\{I + T(I-T)^{-1}\}\underline{U} \\
&= k!\underline{\alpha'}(I-T)^{-k}T^{k-1}\{(I-T)(I-T)^{-1} + T(I-T)^{-1}\}\underline{U} \\
&= k!\underline{\alpha}(I-T)^{-k}T^{k-1}\{(I-T) + T\}(I-T)^{-1}\underline{U} \\
&= k!\underline{\alpha}(I-T)^{-k}T^{k-1}\underline{e}
\end{aligned}
$$

(3.2.3.12)

4 Continuous Time Markov Chain and The Continuous Phase Type Distribution

4.1 Continuous Time Markov Chain

4.1.1 Preliminaries

A Continuous Time Markov Chain $\{X(t) : t \geq 0\}$, also known as Markov Jump, is a Continuous Time Stochastic Process that satisfies the memoryless Markov Property. The current state of the process depends only on the previous state of the process and not any other earlier.

$$\begin{aligned} p_{ij}(0,t) &= p_{ij}(s,s+t) = p_{ij}(t) \\ p_{ij}(s,t) &= prob\{X(t) = j|X(s) = i\} \end{aligned} \tag{4.1.1.1}$$

The chapmann-Kolmogorov Equations are:

For time homogeneous process;

$$p_{ij}(s+t) = \sum_k p_{ik}(s)p_{kj}(t)$$

For time non-homogeneous

$$p_{i,j}(\tau,t) = \sum_k p_{ik}(\tau,s)p_{kj}(s,t) \tag{4.1.1.2}$$

4.1.2 The Generator Matrix and Intensity Rate

Denote the intensity rate from state i to state j as q_{ij}.

$$\text{Define a variable, } \delta_{ij} = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

$$\begin{aligned} q_{ij} &= \lim_{h\to 0} \frac{p_{ij}(h) - \delta_{ij}}{h} \\ &= \begin{cases} \lim_{h\to 0} \frac{p_{ij}}{h} & \text{, for } j \neq i \\ \lim_{h\to 0} \frac{p_{ij}-1}{h} & \text{, for } j = i \end{cases} \end{aligned} \tag{4.1.2.1}$$

Alternatively; $p_{ij} = q_{ij}h + \theta(h)$
Where $\theta(h)$ tends to zero faster than h, i.e.;

$$\begin{aligned}
&\lim_{h\to 0}\frac{\theta(h)}{h} = 0\\
&\therefore \lim_{h\to 0}\frac{p_{ij}}{h} = \lim_{h\to 0}(\frac{q_{ij}}{h} + \frac{\theta(h)}{h})\\
&\qquad\qquad = q_{ij} + \lim_{h\to 0}\frac{\theta(h)}{h} = q_{ij}\\
&\therefore \lim_{h\to 0}\frac{p_{ij}}{h} = q_{ij}
\end{aligned} \tag{4.1.2.2}$$

In a Makov chain, each row of the transition matrix adds up to one, i.e.;

$$\begin{aligned}
\sum_j p_{ij}(h) &= 1\\
\sum_{j\neq i} p_{ij}(h) + p_{ii}(h) &= 1\\
\therefore p_{ii}(h) - 1 &= -\sum_{j\neq i} p_{ij}(h)\\
\lim_{h\to 0}\frac{p_{ii}(h)-1}{h} &= \sum_{j\neq 0}\lim_{h\to 0}\frac{p_{ij}(h)}{h}\\
q_{ij} &= -\sum_{j\neq i} q_{ij}\\
\therefore \sum_{j\neq i} q_{ij} + q_{ii} &= 0\\
\sum_j q_{ij} &= 0
\end{aligned} \tag{4.1.2.3}$$

A generator matrix Q of intensity rates q_{ij} is defined by $Q = ((q_{ij}))$

$$Q = \begin{bmatrix} q_{00} & q_{01} & q_{02} & \cdots & q_{0n}\\ q_{10} & q_{11} & q_{12} & \cdots & q_{1n}\\ q_{20} & q_{21} & q_{22} & \cdots & q_{2n}\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ q_{n0} & q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix} = \begin{bmatrix} -\sum_{j\neq 0} q_{0j} & q_{01} & q_{02} & \cdots & q_{0n}\\ q_{10} & -\sum_{j\neq 1} q_{1j} & q_{12} & \cdots & q_{1n}\\ q_{20} & q_{21} & -\sum_{j\neq 2} q_{2j} & \cdots & q_{2n}\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ q_{n0} & q_{n1} & q_{n2} & \cdots & -\sum_{j\neq n} q_{2j} \end{bmatrix}$$

Rows add up to 0, and elements along the diagonal are negative. For a Markov chain with one absorbing state, the generator matrix can be partitioned into four as follows;

$$Q = \left[\begin{array}{c|cccc} 0 & 0 & 0 & \cdots & 0 \\ \hline q_{10} & -\sum_{j\neq 1} q_{1j} & q_{12} & \cdots & q_{1n} \\ q_{20} & q_{21} & -\sum_{j\neq 2} q_{2j} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ q_{n0} & q_{n1} & q_{n2} & \cdots & -\sum_{j\neq n} q_{2j} \end{array}\right]$$

$$Q = \begin{bmatrix} 0 & \underline{0} \\ t & T \end{bmatrix} \tag{4.1.2.4}$$

4.1.3 Kolmogorov Differential Equations

By Chapman-Kolmogorov equation for time homogeneous case

$$\begin{aligned} p_{ij}(s, s+h) &= p_{ij}(h+t) \\ \therefore p_{ij}(h+h) &= \sum_k p_{ik}(h) p_{kj}(t) \\ &= \sum_{k\neq i} p_{ik}(h) p_{hj}(t) + p_{ii}(h) p_{ij}(t) \end{aligned} \tag{4.1.3.1}$$

$$\begin{aligned} p_{ij}(t+h) - p_{ij}(t) &= \sum_{k\neq i} P_{ij}(h) p_{kj}(t) + p_{ii}(h) p_{ij}(t) - p_{ij}(t) \\ &= \sum_{k\neq i} p_{ik}(h) p_{kj}(t) + \{p_{ii}(h) - 1\} p_{ij}(t) \\ \therefore \frac{p_{t+h} - p_{ij}(t)}{h} &= \sum_{k\neq i} \lim_{h\to 0} \frac{p_{ik}(h)}{h} p_{kj}(t) + \lim_{h\to 0} \{\frac{p_{ii}(h) - 1}{h} p_{ij}(t)\} \\ \therefore p'_{ij}(t) &= \sum_{k\neq i} q_{ik} p_{kj}(t) + q_{ii}(t) \\ &= \sum_k q_{ik} p_{kj}(t) \end{aligned}$$

in matrix form

$$((p'_{ij}(t))) = ((\sum_k q_{ik} p_{kj}(t)))$$

in short form, we can write;

$$P'(t) = QP(t) \tag{4.1.3.2}$$

Solutions to Kolmogorov Equations

let $P(t) = e^{Qt}$ which is called exponential matrix.

$$P(t) = \sum_{k=0}^{\infty} \frac{(Qt)^k}{k!} = \sum_{k=0}^{\infty} \frac{Q^k t^k}{k!} = \sum_{k=0}^{\infty} \frac{t^k Q^k}{k!} \tag{4.1.3.3}$$

differentiating both sides of the equation with respect to t

$$\frac{d}{dt}P(t) = \sum_{k=0}^{\infty} \frac{kt^{k-1}Q^k}{k!} = \sum_{k=1}^{\infty} \frac{kt^{k-1}Q^k}{k!} = \sum_{k=1}^{\infty} \frac{kt^{k-1}QQ^{k-1}}{k!}$$

$$\frac{d}{dt}P(t) = Q\sum_{k=1}^{\infty} \frac{kt^{k-1}Q^{k-1}}{k!} = Q\sum_{k=1}^{\infty} \frac{(Qt)^{k-1}}{(k-1)!} = Qe^{tQ} = Qe^{Qt} = QP(t) \tag{4.1.3.4}$$

$\therefore P'(t) = QP(t)$ and $P(t) = e^{Qt}$ are solutions to the Kolmogorov backward equation.
Or;

$$\frac{d}{dt}P(t) = \sum_{k=1}^{\infty} \frac{t^{k-1}Q^{k-1}}{(k-1)!}Q = [\sum_{k=1}^{\infty} \frac{t^{k-1}Q^{k-1}}{(k-1)!}]Q = e^{tQ}Q = e^{Qt}Q$$

$$\therefore P'(t) = P(t)Q \tag{4.1.3.5}$$

Conclusion: The exponential matrix; $P(t) = e^{Qt}$ is a solution to both the forward and backward Kolmogorov Differential Equations.

Eigenvalues and Eigenvectors

To determine Eigenvalues of Q we solve the equation $|Q - \lambda I| = 0$
If the Eigenvalues are all distinct then Q can be expressed as $Q = UDU^{-1}$; Where D is a diagonal matrix whose elements are the Eigenvalues, U is a matrix of right eigen vectors and U^{-1} is the inverse of U.

$$Q^2 = QQ = (UDU^{-1})(UDU^{-1})$$
$$= UDU^{-1}UDU^{-1}$$

but $U^{-1}U = I$, therefore we have;

$$Q^2 = UDIDU^{-1} = UD^2U^{-1} \tag{4.1.3.6}$$

Similarly, for the third order we have;

$$Q^3 = QQ^2 = (UDU^{-1})(UD^2U^{-1})$$
$$= UDU^{-1}UD^2U^{-1}$$

but $U^{-1}U = I$, therefore we have;

$$Q^3 = UDID^2U^{-1} = UD^3U^{-1} \tag{4.1.3.7}$$

In general, if a square matrix Q can be diagonized, then a function of Q is just the same as multiplying its eigenvalues matrix by the function of the diagonal matrix and then multiplying the result by the inverse of the matrix of the eigenvalues.

$$\therefore Q^k = UD^kU^{-1} \tag{4.1.3.8}$$

To determine U we solve for $\underline{x}$ in the equation

$$Q\underline{x} = \lambda \underline{x}$$

$$\begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1m} \\ q_{21} & q_{22} & \cdots & q_{2m} \\ \vdots & & & \\ q_{m1} & q_{m2} & \cdots & q_{mm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \lambda_j \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad ; j = 1, 2, 3, \ldots \tag{4.1.3.9}$$

We suppose we have m distinct eigenvalues.

$$\begin{aligned} \therefore P(t) &= \sum_{k=0}^{\infty} \frac{t^k Q^k}{k!} \\ &= \sum_{k=0}^{\infty} t^k \frac{(UDU^{-1})^k}{k!} \\ &= \sum_{k=0} \infty \frac{t^k UD^k U^{-1}}{k!} \end{aligned} \tag{4.1.3.10}$$

since U is constant with respect to k it can therefore be factored out of the brackets and we have;

$$\therefore P(t) = U[\sum_{k=0}^{\infty} \frac{(tD)^k}{k!}]U^{-1}$$
$$= Ue^{tD}U^{-1}$$

But

$$\sum_{k=0}^{\infty} \frac{t^k D^k}{k!} = \sum_{k=0} k \frac{t^k}{k!}$$

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots \\ 0 & 0 & 0 & \vdots & \lambda_m \end{pmatrix} = \begin{pmatrix} \frac{\sum_{k=0}^{\infty} t^k \lambda_1^k}{k!} & 0 & 0 & \dots & 0 \\ 0 & \frac{\sum_{k=0}^{\infty} t^k \lambda_2^k}{k!} & 0 & \dots & 0 \\ \vdots \\ 0 & 0 & 0 & \dots & \frac{\sum_{k=0}^{\infty} t^k \lambda_3^k}{k!} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{\lambda_1 t^k}{k!} & 0 & 0 & \dots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{\lambda_2 t^k}{k!} & 0 & \dots & 0 \\ \vdots \\ 0 & 0 & 0 & \dots & \sum_{k=0}^{\infty} \frac{\lambda_m t^k}{k!} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1^t & 0 & 0 & \dots & 0 \\ 0 & \lambda_2^t & 0 & \dots & 0 \\ \vdots \\ 0 & 0 & 0 & \dots & \lambda_m^t \end{pmatrix}$$

$$\therefore P(t) = U \sum_{k=0}^{\infty} \frac{(tD)^k}{k!} U^{-1}$$

becomes

$$P(t) = U \begin{pmatrix} e^{\lambda_1^t} & 0 & 0 & \dots & 0 \\ 0 & e^{\lambda_2^t} & 0 & \dots & 0 \\ \vdots \\ 0 & 0 & 0 & \dots & \lambda_m^t \end{pmatrix} U^{-1} \quad \textbf{(4.1.3.11)}$$

4.2 Continuous Phase-Type (PH) Distribution

4.2.1 Preliminaries

A continuous phase type distribution describes the distribution of time that a process will take to reach the absorbing state. The model requires that the process starts in a transient state and that there is one absorbing state. This chapter discusses the fundamental properties of a continuous phase-type distribution. We shall derive the probability density function and the cumulative density function of a continuous Phase-type distribution. We shall also derive the moment generating function and the probability generating function of a phase type distribution that will be used to determine the expected time that the process will reach its absorbing state and also the variance of the phase type distribution.

Consider a continuous time Markov process with m transient states $\{1,2,\ldots,m\}$ and one absorbing state $m+1$. From Equation (4.1.2.4), the infinitesimal generator matrix Q can be represented by a 2×2 block matrix

$$Q = \begin{bmatrix} \underline{O} & 0 \\ T & t \end{bmatrix}$$

Where:

- T is an $m\times m$ matrix of intensity rates from one transient state to another transient state.
- t is an $m\times 1$ matrix of intensity rates from a transient state to the absorbing state.
- $\underline{O}$ is $1\times m$ vector of zeros

Let $\underline{e}$ be an $m\times 1$ column vector consisting of 1. Since rows of a generator matrix, Q add up to 1, Recall the column matrix $\underline{e}$ of 1's from equation (3.1.2.2);

$$\therefore T\underline{e}+t = 0$$
$$t = -T\underline{e}$$

Therefore, the infinitesimal generator matrix Q can partitioned as;

$$Q = \begin{bmatrix} T & -T\underline{e} \\ \underline{O} & 0 \end{bmatrix} \quad \textbf{(4.2.1.1)}$$

Let $\underline{\beta} = \{\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}\}$ denote the vector of probabilities that the process starts in states $1, 2, \ldots, m,$ and $m+1$ respectively. Since a phase type distribution does not start in a transient state, therefore the probability that the probability that the process starts in the absorbing state $m+1$ is zero, hence, $\alpha_{m+1} = 0.$ The initial probability vector can therefore be rewritten as $\underline{\beta} = \underline{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}.$

Let Z be the random variable denoting the time taken by the process to reach the absorbing state such that $Z \quad Ph(\underline{\alpha}, T).$

$\therefore f(z) =$ the probability that the process will be absorbed at time z.

$=$ the probability of starting in any transient state, being in any transient state at time z-h and being in the absorbing state at time z where h->0.

(4.2.1.2)

$1 - F(z) =$ the probability that the process will not have reached the absorbing state at time z

$$= Prob\{Z > z\} = \underline{\alpha} e^{Tz} \underline{e}$$

(4.2.1.3)

$F(Z) =$ the cumulative distribution of the phase type distribution.

$=$ the probability that the process will be in the absorbing state before or at time z.

$$= Prob\{Z \leq z\} = 1 - \underline{\alpha} e^{Tz} \underline{e}$$

(4.2.1.4)

Theorem 4.2.1. *The probability density function of a random variable Z that follows a phase type distribution, i.e.; z- $Ph\{\underline{\alpha}, T\}$, is given by;*

$$f(z) = -\underline{\alpha} e^{Tz} T \underline{e} \qquad \textbf{(4.2.1.5)}$$

Proof. The density function $f(z)$ can be determined by differentiating the cumulative function $F(z)$ with respect to the random variable z as follows.

$$\begin{aligned}
\therefore f(z) =& \frac{dF(z)}{dz} \\
=& \frac{d}{dz}(1 - \underline{\alpha} e^{TZ} \underline{e}) \\
=& \frac{d}{dz}(-\underline{\alpha} e^{TZ} \underline{e}) \\
=& \frac{d}{dz} - \underline{\alpha}(e^{TZ} \underline{e}) \\
=& -\underline{\alpha} T e^{TZ} \underline{e} \\
=& -\underline{\alpha} e^{TZ} T \underline{e} \\
=& \underline{\alpha} e^{TZ} t
\end{aligned}$$

(4.2.1.6)

□

4.2.2 The Moment Generating Function of a Continuous Phase Type Distribution

Theorem 4.2.2. *If Z is a random variable that follows a phase-type distribution with initial probability $\underline{\alpha}$, then the moment generating function of the continuous phase type distribution defined as $Ph(\underline{\alpha},T)$ is given by;*

$$E(z^n) = (-1)^n n! \underline{\alpha} T^{-n} \underline{e} \tag{4.2.2.1}$$

Proof.
Let $L_z(s)$ denote the Laplace transform of a phase type distribution $Z\text{-}Ph(\underline{\alpha},T)$

$$F(s) = L\{f(t)\} \tag{4.2.2.2}$$

$$\begin{aligned} L_z(s) = E(e^(sz)) &= \int_0^\infty e^{-sz} f(z) dz \\ &= \int_0^\infty e^{-sz} \underline{\alpha} e^{Tz} t dz \\ &= \underline{\alpha} \int_0^\infty e^{-sz} e^{Tz} t dz \\ &= \underline{\alpha} \int_0^\infty e^{(-sz+Tz)} dzt \\ &= \underline{\alpha} (\int_0^\infty e^{-(sI-T)z} dz) t \\ &= -\underline{\alpha}(sI-T)^{-1} [e^{-(sI-T)z}]_{z=0}^\infty t \\ &= -\underline{\alpha}(sI-T)^{-1}[0-1]t \\ &= \underline{\alpha}(sI-T)^{-1} t \end{aligned} \tag{4.2.2.3}$$

$$\begin{aligned} E(z^n) &= (-1) \frac{d^n}{dz^n} L_z(s)]_{s=0} \\ \therefore E(z^n) &= (-1)^n \frac{d^n}{dz^n} [\underline{\alpha}(sI-T)^{-1} t]_{s=0} \\ &= (-1)^n \underline{\alpha} \frac{d^n}{dz^n} [(sI-T)^{-1}]_{s=0} t \end{aligned} \tag{4.2.2.4}$$

Consider;

$$\frac{d^n}{dz^n}(sI-T)^{-1} \tag{4.2.2.5}$$

When $n = 1$

$$\frac{d}{ds}(sI-T)^{-1} = (-1)(sI-T)^{-2} \tag{4.2.2.6}$$

When $n = 2$

$$\begin{aligned}\frac{d^2}{ds^2}(sI-T)^{-1} &= \frac{d}{ds}\left[\frac{d}{ds}(sI-T)^{-1}\right]\\ &= \frac{d}{ds}\left[(-1)(sI-T)^{-2}\right]\\ &= (-1)\frac{d}{ds}\left[(sI-T)^{-2}\right]\\ &= (-1)(-2)\left[(sI-T)^{-3}\right]\\ &= (-1)^2(2)\left[(sI-T)^{-3}\right]\end{aligned} \tag{4.2.2.7}$$

When $n = 3$

$$\begin{aligned}\frac{d^3}{ds^3}(sI-T)^{-1} &= \frac{d}{ds}\left[\frac{d^2}{ds^2}(sI-T)^{-1}\right]\\ &= \frac{d}{ds}\left[(-1)^2(2)(sI-T)^{-3}\right]\\ &= (-1)^2(2)\frac{d}{ds}\left[(sI-T)^{-3}\right]\\ &= (-1)^2(-2)(2)(-3)\left[(sI-T)^{-4}\right]\\ &= (-1)^3(2)(3)\left[(sI-T)^{-4}\right]\\ &= (-1)^3 3!\left[(sI-T)^{-4}\right]\end{aligned} \tag{4.2.2.8}$$

Therefore, by induction;

$$\frac{d^n}{dz^n}(sI-T)^{-1} = (-1)^n n!(sI-T)^{-(n+1)} \tag{4.2.2.9}$$

The moment generating function is therefore determined as follows;

$$\begin{aligned}E(z^n) &= \underline{\alpha}(-1)^n(-1)^n n!\left[(sI-T)^{-(n+1)}\right]_{s=0} t\\ &= \underline{\alpha}(-1)^{(n+n)} n!\left[(sI-T)^{-(n+1)}\right]_{s=0} t\\ &= \underline{\alpha}(-1)^{(n+n)} n!(-T)^{-(n+1)} t\\ &\text{but } t = -T\underline{e}\text{ ; from equation (4.2.1.1)}\\ \therefore E(z^n) &= \underline{\alpha}(-1)^{(n+n)} n!(-1)^{-(n+1)}(T)^{-(n+1)}(-T\underline{e})\\ &= \underline{\alpha}n!(-1)^{(n+n)} n!(-1)^{-(n+1)}T^{-(n+1)}(-1)(T\underline{e})\\ &= \underline{\alpha}n!(-1)^{(n+n-(n+1)+1)}T^{-(n+1)+1}\underline{e}\\ &= \underline{\alpha}n!(-1)^n T^{-n}\underline{e}\\ &= (-1)^n\underline{\alpha}n!T^{-n}\underline{e}\end{aligned} \tag{4.2.2.10}$$

The expected time to absorption is the first moment and is given as:

$$E(z) = (-1)^1 \underline{\alpha} 1! T^{-1} \underline{e}$$
$$E(z) = -\underline{\alpha} T^{-1} \underline{e} \tag{4.2.2.11}$$

The second moment is :

$$E(z^2) = (-1)^2 \underline{\alpha} 2! T^{-2} \underline{e}$$
$$E(z^2) = 2\underline{\alpha} T^{-2} \underline{e} \tag{4.2.2.12}$$

Hence, the variance of a phase type random variable is;

$$\begin{aligned} Var(z) &= E(z^2) - \{E(z)\}^2 \\ &= (2\underline{\alpha} T^{-2} \underline{e}) - (-\underline{\alpha} T^{-1} \underline{e})^2 \\ &= (2\underline{\alpha} T^{-2} \underline{e}) - (\underline{\alpha} T^{-1} \underline{e})^2 \end{aligned} \tag{4.2.2.13}$$

5 Application

5.1 preliminaries

This chapter applies the Continuous Phase type distribution to actuarial science. Actuarial functions including life annuities and life assurance were estimated using the continuous phase type distribution. These two functions were then used to determine premiums by solving the equation of value. Data of AM92 of the Actuarial Orange tables was used in this study(Tables, 1980). This data was considered since the orange tables are frequently used for Actuarial Examinations as well as Actuarial Science courses. Contributions to Actuarial study such as (Bowers & of Actuaries, 1986) and (Scott, 1999) also makes reference to actuarial tables. Life table functions from AM92 was used to estimate parameters of the phase-type distribution used in this study (Tables, 1980).

An introduction to Physiological Aging, phase-type law of mortality and Parameters estimation of the phase type distribution were determined in the first section of this chapter. Actuarial present value of Life Assurance and Life Annuities will be fitted and net premiums payable to whole life insurance policies will be determined in the next sections. This chapter will conclude by testing the goodness of fit of the model through comparing the fitted actuarial functions to those estimated directly from life-table functions.

5.2 Physiological Aging and Parameter Estimation

Mortality rates are observed to increase with age. Age specific morality increases with age. Consequently, the probability of survival decreases with age. The observed one year probability of death and survival were plotted in figure 1 below. The population's initial age was 17. The maximum attainable lifespan in this population was 120. At this age, all lives die. The probability of survival beyond 120 years is zero. Hence, the probability of death is 1.

The data shows low mortality rates at low ages. The One year probability of death curve is fairly a straight line with a slope close to zero from ages 17 to 50. The slope gradually increases between ages 45 to 80. The rate of death increases at a higher rate from age 80. This is shown by the higher slope from age 80 (see Figure 1).

However, in real world phenomenon, there is no specific maximum lifespan. There are cases where lives survive beyond the expected "limiting age". The famous Lee-Carter model (Lee & Carter, 1992) and the CBD model (Cairnsab, Blakec, & Dowdd, 2005) suggest that in the future lives will survive to infinity. It is therefore important to model mortality by considering that the limiting age is infinity. The future lifetime random variable of the phase-type distribution takes values from 0 to infinity.

This study makes use of the idea of physiological aging. The idea of aging process was first introduced by (Jones, 1956).He argued that aging is a process under which the body

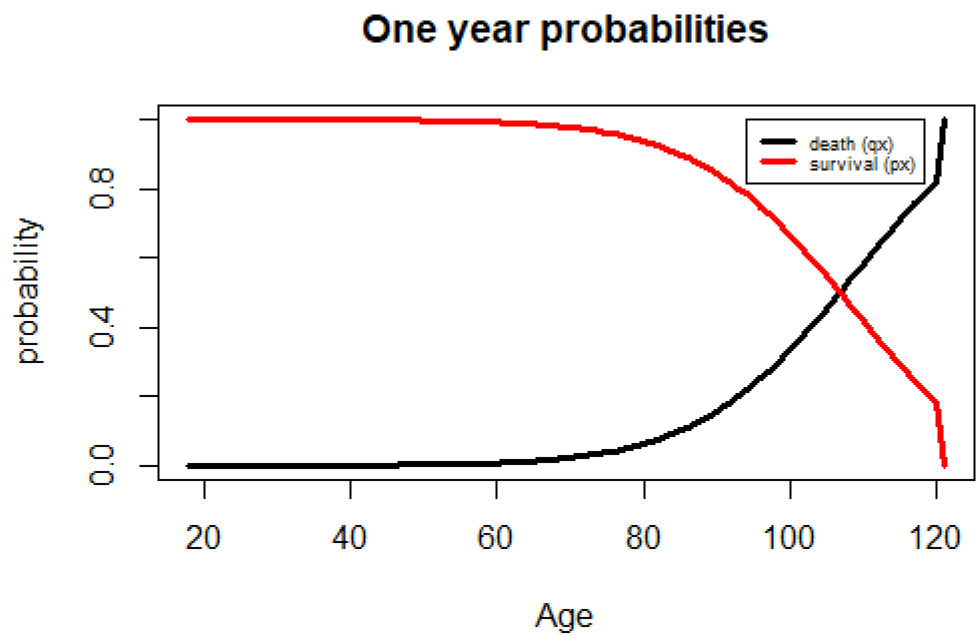


Figure 1. Observed one year probability of death and survival

undergoes biological and physical changes. These changes are irreversible and can be slower at lower age while aging is rapid at higher age-specificity. Aging is associated with deterioration in health of individual (or age group) implying that older individuals are comparatively more susceptible to diseases and consequently at higher risk to death than younger individuals (Jones, 1959). This aging phenomenon was termed *physiological aging* by Dr Lin and Ms Liu (Lin & Liu, 2007).

Physiological age is a hypothetical variable that explains a significant changes in the biological composition of human anatomy. This study refers to the former as *phases*. Death in older ages is highly linked to the degeneration in the health status of individuals which is largely explained by the languishing of important body organs, tissues and general immune system (Sehl & Yates, 2001).

An individual in phase (i) can either move to the next phase (i+1) or die. Movement from one phase to the next follow an exponential distribution with parameter λ_i,hence, the mean is $(1/\lambda_i)$. The transition intensity from phase (i) to phase (i+1) is therefore λ_i. Previous studies proved that there is a slow average change in degeneration of health status. Thus, the transition rate from any phase to the next is constant.

$$q_{i,i+1} = \lambda_i = \lambda, \text{ for } i = 1, 2, ..n \text{ where } n \text{ is the number of phases.} \qquad \textbf{(5.2.0.1)}$$

A life cannot move to a phase that does not come exactly after it. That is $q_{ij} = 0$ if $j \geq 2$ Death is the one absorbing state in this model and is caused either by accident or death due to health complications represented by physical aging. Besides, aging is a continuous process and time to death is a continuous random variable, therefore, the future life time of a life follows a Phase-Type distribution.Denote the rate of death of individuals in phase (i) by $q_{i,0} = q_i$.

$\therefore q_i = a + b \times i^c$ Where a is component of death rate due to accident and $b \times i^c$ is due to the aging process.

Recall: Rows of a generator matrix add up to one (see equation **(4.2.1.1)**).

$$\begin{aligned} \therefore \sum_j q_{ij} &= 0 \\ q_{ii} &= -\sum_{j \neq i} q_i j \\ q_{ii} &= -(q_{i,0} + q_{i,i+1} + q_{i,i+2} + q_{i,i+3} + ...) \\ q_{ii} &= -(q_i + \lambda + 0 + 0 + \ldots) \\ q_{ii} &= -(q_i + \lambda) \end{aligned} \tag{5.2.0.2}$$

The resulting generator matrix is of coxian nature. More on Coxian representation and other Phase-Type distributions will not be covered in this paper but are available in other literatues (Asmussen, Nerman, & Olsson, 1996) and (He & Zhang, 2007) is represented by the transitions rates matrix T below:

$$Q = \left[\begin{array}{c|cccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline q_1 & -(\lambda + q_1) & (\lambda) & 0 & 0 & \ldots & 0 \\ q_2 & 0 & -(\lambda + q_2) & (\lambda) & 0 & \ldots & 0 \\ q_3 & 0 & 0 & -(\lambda + q_3) & (\lambda) & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ q_m & 0 & 0 & 0 & 0 & \ldots & -qm \end{array}\right] \tag{5.2.0.3}$$

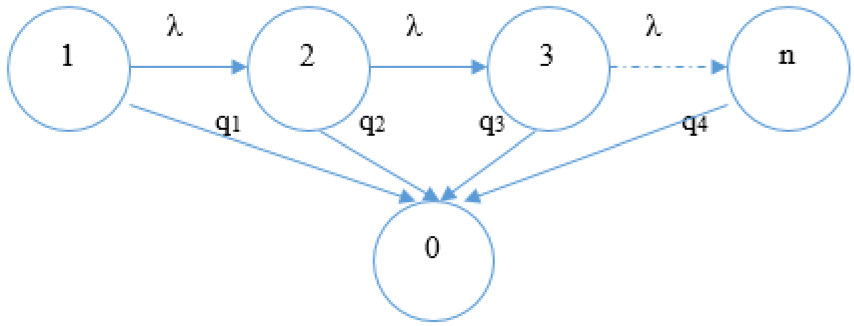


Figure 2. Phase-type Law of Mortality

λ	**a**	**b**	**c**	MSE
3E0	1.20E-03	1.13E-10	4.2E0	1.27E-06

Table 1. Estimated Parameters of Phase-Type Law of Mortality

In the case of AM92, the life starts at age 17, hence, the initial probability vector is $\alpha = (1,0,...,0)$. Figure 2 is a graphical representation of the Phase-type law of mortality of coxian nature. Phase 0 denotes the death phase.

One year mortality rates curve is fairly smooth and increasing. It is therefore sufficient to have a single estimation period. This is contrary to Lin and Liu's study in 2007, where they had to estimate parameters of the transition matrix due to infant mortality, accident hump and mortality due to aging (Lin & Liu, 2007).

The four parameters $(\lambda, a, b,$ and $c)$ are estimated using the Nelder-Mead simplex algorithm method. The objective is to minimize the mean squared error for the probability of future lifetime(Nelder & Mead, 1965). The functions $s_x.q_x$ were determined for each age $x = 17, 18, .., 120$

Let $\widehat{s_X.q_x}$ denote the fitted probability that the future life time of a life in phase 1 will die after x years. therefore,

$$\widehat{s_X.q_x} = prob[\text{Future lifetime is x}]$$
$$\widehat{s_X.q_x} = f(x \leq Z < x+h); \qquad h->0 \tag{5.2.0.4}$$

Where $z - PH(T, \alpha)$
objective:
minimize $MSE = \frac{1}{n}\sum_x (s_x.q_x - s_X\hat{.}q_x)^2$

Table 1 shows the estimated parameters and the resulting mean squared error. The MSE is relatively small. Besides, the fitted survival function fairly follows the observed survival (figure 3). Thus, the estimated parameters are good for modelling the law of mortality of phase-type.

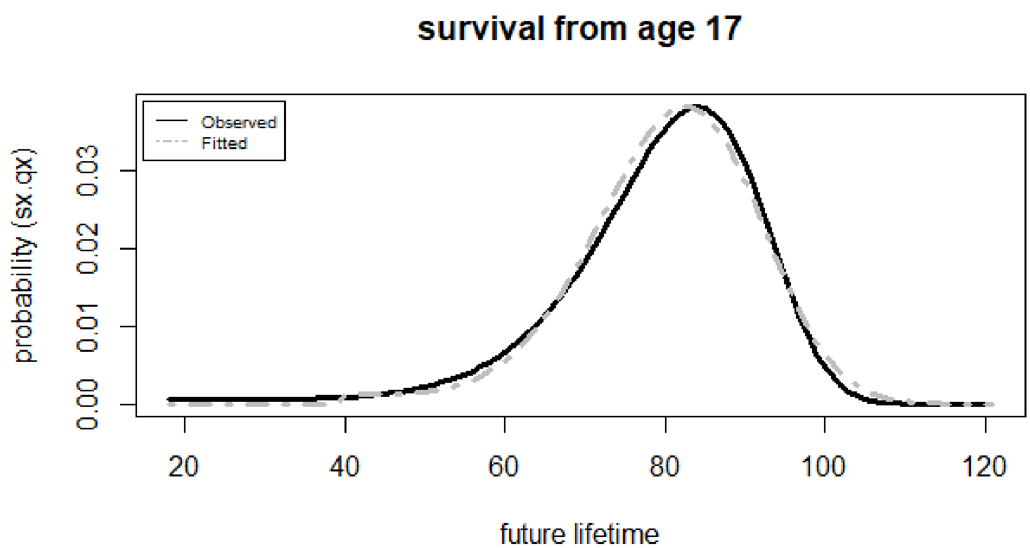


Figure 3. Heterogeneous Distribution

5.3 Actuarial Functions

This section evaluates assurances, annuities and premiums payable for a whole life assurance contract. Consider whole life insurance policy to a life aged $x = 17, 18, \ldots, 200$ paying a sum assured of Ksh. 1 immediately on the death of the policy holder. Annual premiums are payable through out the life of the policyholder and there are no expenses charged on the policy.

Actuarial Assurance

An assurance is the expected present value of the benefit. Since the a benefit of 1 is payable immediately on death of the assured life, the assurance is the expected value of Ksh. 1 discounted for z years where z is the future lifetime of the assured life at a force of interest σ.

$$Assurance = A_x = E[e^{-\sigma * z}]$$

$$but L_z(s) = E(e^{sz}) = \underline{\alpha}(sI - T)^{-1}t$$

$\therefore$ by induction

$$A_x = \underline{\alpha}(-\sigma I - T)^{-1}t \qquad \textbf{(5.3.0.1)}$$

Given an annual interest rate i, the force of interest σ is estimated as $\sigma = \ln(1+i)$. There is a large deviation between assurances fitted by Phase type model and the non-stochastic model at lower ages (between 27 to 97 years) (see Figure 4). Deviations are smaller for younger and older ages for the two interest rate levels and are even smaller when interest rate is higher at 6% than at 4%.

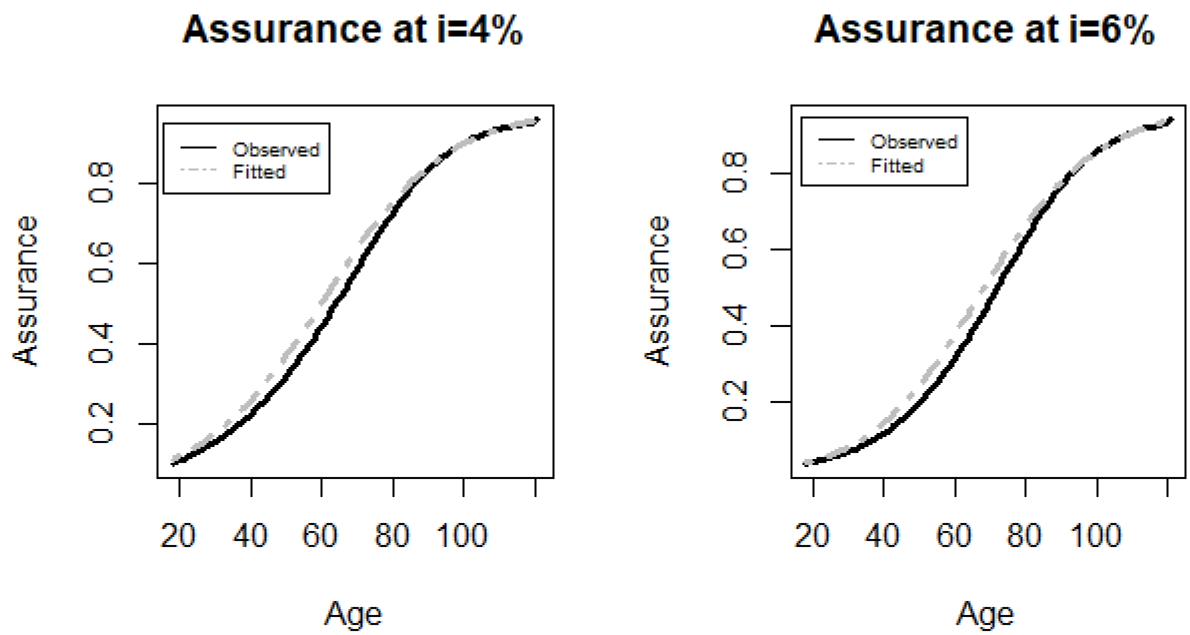


Figure 4. Assurance at 4% and 6% interest rate per annum

Actuarial Annuity

Actuarial Annuities are expected present value of regular payments. Consider level annuity premiums of Ksh. 1 paid for the whole life of the assured life. Premiums are paid at the beginning of the year such that the Actuarial Present Value of annuity for a life aged x, denoted as a_x is given as: $a_x = E[a_z]$

where z is the time to death of a life aged x and Z follows the phase-type distribution.

$$
\begin{aligned}
a_x &= E[\frac{(1-v^z)}{d}] \\
&= \frac{1-E[V^z]}{d} \\
&= \frac{(1-A_x)}{d} \\
\text{Where} d &= \frac{i}{1+i}
\end{aligned}
\quad (5.3.0.2)
$$

Present value of annuities were estimated from present value of assurances at interest of 4% and 6% (Figure 5). Model's estimates fairly follow the shape of observed expected life annuities. Life annuities decrease with age. The values are higher when interest rate is 4% compared to 6%. Just as the assurances, annuity deviations for both interest rates are lower for younger and older ages as compared to the middle ages.

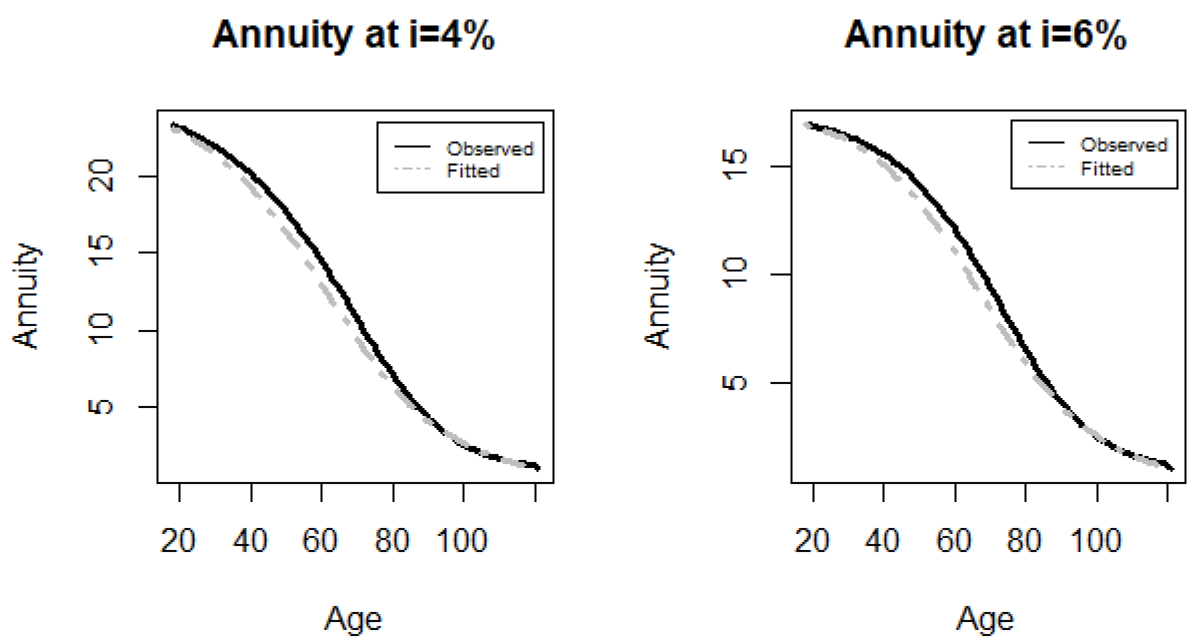


Figure 5. Annuity at 4% and 6% interest rate per annum

Premiums

Consider a life insurance policy paying a sum assurance of Ksh.1 immediately on death of the policy holder. Level premiums are payable annually. For simplicity, assume that there is no surrender and no expenses are incurred. Level premiums are determined from the equation of value. Let P_x be the annual premium payable for such policy.

$$\therefore P_x a_x = A_x$$
$$P_x = \frac{A_x}{a_X} \quad \textbf{(5.3.0.3)}$$

There are negligible errors between observed and Phase-Type fitted premiums for both 4% and 6% interest rates in the sense that there is minimal deviations compared to assurance and annuity curves, premiums curves Fairly fall on each other as seen in figure 6.

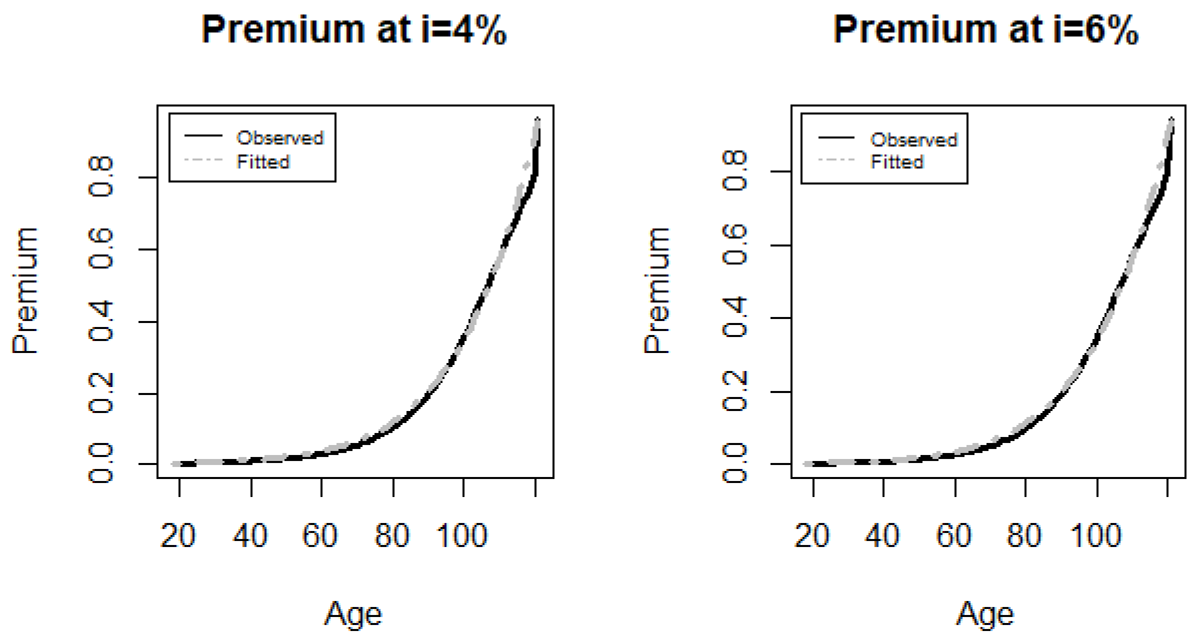


Figure 6. Premiums at 4% and 6% interest rate par annum

5.4 Goodness of Fit

This section discusses the goodness of fit for each actuarial function discussed in earlier sections of this chapter at 4% and 6% by comparing functions from Phase-Type distributions with functions from life tables, therein determining the mean squared error (MSE) and mean absolute error (MAD).

$$MAD = \frac{1}{n}\sum_{i=1}^{n} |Observed - fitted|$$
$$MSE = \frac{1}{n}\sum_{i=1}^{n} (Observed - fitted)^2$$

Table 2 is a summary of the model's goodness. Entries in the first row including Ass.,Ann., and Prem. are abbreviations for Assurance, Annuity and Premium. MAD and MSE for survival function are significantly low. This indicate that the Phase-Type physiological aging process is a good fit for modelling the future lifetime of a population Similarly for Actuarial functions including Assurance, Annuity and Premium, the errors are significantly low. Of the actuarial functions, Premiums have the lowest errors, followed by Assurance functions. Though Annuities have the highest deviations, their mean errors are still significantly low, hence,are a good fit.

Error	Survival	Ass. (4%)	Ass. (6%)	Ann. (4%)	Ann. (6%)	Prem. (4%)	Prem. (6%)
MAD	0.003	0.003	0.003	0.066	0.051	0.001	0.001
MSE	0.000	0.001	0.001	0.072	0.036	0.000	0.000

Table 2. Test of Goodness of Fit

6 Conclusion

6.1 Summary of Findings

Chapters 3 and 4 elaborated the phase type distribution by deriving the distribution and moments of the distribution. These were the basis of the application in chapter 5 of the study. Results showed that one estimation period was enough for the hypothetical life table of AM92. The transition intensity to the next age is constant through out the years hence, the time taken in each phase follow a Poisson distribution with a constant Poisson parameter λ. The following is a short summary of discrete and continuous distribution and their moments.

Discrete Time

$$\begin{aligned} f(z) &= Prob\{Z=z\} = \underline{\alpha}' T^z \underline{U} \\ F(z) &= prob\{Z \leq z\} = 1 - \underline{\alpha}' T^z \\ G(s) &= E\{s^z\} = \alpha_0 + \underline{\alpha}' s(I-sT)^{-1}\underline{U} \\ A_x &= Ev^z = \alpha_0 + \underline{\alpha}' v(I-vT)^{-1}\underline{U} \\ a_x &= Ea_{\overline{n}|} = E\left[\frac{1-V^z}{d}\right] = \frac{1-A_x}{d} \end{aligned} \tag{6.1.0.1}$$

Continuous Time

$$\begin{aligned} f(z) &= Prob\{z \leq Z < z+h\} = -\underline{\alpha}' e^{Tz} T \underline{e} == \underline{\alpha}' e^{Tz} t \\ F(z) &= prob\{Z \leq z\} = 1 - \underline{\alpha}' e^{Tz} \underline{e} \\ M_z(n) &= E\{e^{sz}\} = \underline{\alpha}(sI-T)^{-1} t \\ \bar{A}_x &= Ee^{-\sigma z} = \underline{\alpha}(-\sigma I - T)^{-1} t \\ \bar{a}_x &= E\bar{a}_{\overline{n}|} = E\left[\frac{1-e^{-\sigma z}}{\sigma}\right] = \frac{1-\bar{A}_x}{\sigma} \end{aligned} \tag{6.1.0.2}$$

6.2 Conclusion

Results showed that phase-type distribution of coxian nature is applicable in the computation of actuarial functions. The study of the model's goodness of fit through the mean absolute error and mean squared error shows good approximation of functions computed directly from life tables functions and those determined after phase-type modelling. Results showed that there is small margin of deviation between premiums computed directly from life table functions and those computed by applying a Phase-Type distribution. Also, there is a large underestimation of assurances and overestimation of annuities between the ages40and90. However, there are small deviations for whole life net premiums

6.3 Recommendation

This study has proven that Phase-Type distribution is the most appropriate model for life insurance premium computations and thus Life Insurance Companies and Pension providers should adopt the model for product pricing. This model should also be included in actuarial science study curriculum to equip students with variety of distributions for actuarial functions computations.

6.4 Future Research

Hypothetical AM 92 life table was used in this study for model building. This study can be extended to real life mortality data. Also, the study was limited to computation of mean present value of whole life assurance, annuities and net premiums. A further research can be done to other life insurance contracts such as endowment, pure endowment and term insurance contracts. Also, other functions including reserves can be considered.

6.5 limitations

The phase-type law of mortality is highly dependant on life table functions in that parameters of the transition matrix and generator matrix are solely determined by considering life table probability of death and survival. Besides, at present, there is no available data of the hypothetical physiological aging process model is built on which the phase-type law of mortality is based in that the physiological aging probabilities have to be estimated before the phase-type model is built.

Bibliography

Asmussen, S. (1989). Risk theory in a markovian environment. *Scandinavian Actuarial Journal*, *1989*(2), 69–100.

Asmussen, S., Nerman, O., & Olsson, M. (1996). Fitting phase-type distributions via the em algorithm. *Scandinavian Journal of Statistics*, 419–441.

Bowers, N., & of Actuaries, S. (1986). *Actuarial mathematics*. Society of Actuaries. Retrieved from https://books.google.co.ke/books?id=jrjuAAAAMAAJ

Cairnsab, A. J., Blakec, D., & Dowdd, K. (2005). A two-factor model for stochastic mortality with parameter uncertainty.

de Moivre, A. (1725). *Annuities upon lives, or, the valuation of annuities upon any number of lives, as also, of reversions: To which is added, an appendix concerning the expectations of life, and probabilities of survivorship.* London printed, and Dublin re-printed, by and for S. Fuller.

Gompertz, B. (1825). Xxiv. on the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq. frs &c. *Philosophical transactions of the Royal Society of London*(115), 513–583.

Graunt, J. (1973). 1662. *Natural and Political Observations Made Upon The Bills of Mortality London*.

He, Q.-M., & Zhang, H. (2007). Coxian approximations of matrix-exponential distributions. *Calcolo*, *44*(4), 235–264.

Jones, H. B. (1956). A special consideration of the aging process, disease, and life expectancy. *Advances in biological and medical physics*, *4*, 281–337.

Jones, H. B. (1959). The relation of human health to age, place and time. *Handbook of Aging and the Individual*, 336–63.

Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting us mortality. *Journal of the American statistical association*, *87*(419), 659–671.

Lin, X. S., & Liu, X. (2007). Markov aging process and phase-type law of mortality. *North American Actuarial Journal*, *11*(4), 92–109.

Makeham, W. M. (1860). On the law of mortality and the construction of annuity tables. *Journal of the Institute of Actuaries*, *8*(6), 301–310.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The computer journal*, *7*(4), 308–313.

Organization, W. H. (2015). *World health statistics 2016: monitoring health for the sdgs sustainable development goals.* Author.

Organization, W. H. (2020). *World health statistics 2016: monitoring health for the sdgs sustainable development goals.* Author.

Scott, W. F. (1999). Life assurance mathematics.

Sehl, M. E., & Yates, F. E. (2001). Kinetics of human aging: I. rates of senescence between ages 30 and 70 years in healthy people. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *56*(5), B198–B208.

Tables, G. (1980). Formulae and tables for actuarial examinations. *Institute Of Actuaries and Faculty Of Actuaries*.