UNIVERSITY OF NAIROBI

Phase Type Models Applied in Estimation of Aggregate Claim Losses of Secondary Cancer Cases

BY

Cynthia Mwende Mwau

I80/57255/2020

PhD Thesis Submitted to the Department of Mathematics, Faculty of Science and Technology, University of Nairobi in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy in Actuarial science of the University of Nairobi

November 2022

Declaration and Approval

I Cynthia Mwende Mwau declare that this is my original work and has never been presented for any academic award in any other institution.

	24/11/2022
Signature	Date

CYNTHIA MWENDE MWAU

I80/57255/2020

In my capacity as a supervisor of Cynthia Mwende Mwau, I certify that this PhD Thesis has my approval for submission.

	24 November 2022
Signature	Date

Prof. P.G.O Weke
Department of Mathematics, University of Nairobi.

	24th November 2022
Signature	Date

Dr Davis Bundi
Department of Mathematics, University of Nairobi.

	24/11/2022
Signature	Date

Prof. J. A. M Otieno
Department of Mathematics, University of Nairobi.

Abstract

Aggregate losses can be applied widely in areas of actuarial science as well as financial mathematics. They can be calculated using the collective risk model which sums random losses involving both claim severity and claim frequency. Impact of claim severity on aggregate losses has been well explored in previous research while less research has been done on impact of claim frequency on aggregate losses especially using phase type distributions which motivates this study.

In this research we improve on calculation of aggregate losses by introducing phase type distributions in modeling claim frequency, construct phase type Poisson Lindley, determine their properties and parameter estimation. This research also determines how to get matrix parameters of phase type distributions, construct phase type compound probability generating function and apply the proposed models to secondary cancer cases in Kenya to demonstrate their advantage. Phase type distributions have one of their parameter as a matrix hence they can be used to model claim frequency for diseases which have multiple stages of transition and data which applies bonus malus system. The phase type distributions considered in this research are Panjer class $(a, b, 0)$, class $(a, b, 1)$ and Poisson Lindley distributions. Matrices calculated using Chapman-Kolmogorov equation have shown to fit well in the phase type distributions. The concept of survival analysis (Kaplan-Meier) is used to estimate the transition probabilities of the matrix parameters and the long run probabilities represent the row vector Υ. Severity distributions considered are one and two parameter Poisson Lindley distribution, Pareto, Generalized Pareto and Wei-bull distributions. Method of moments is used in estimation of parameters of the severity distributions while Panjer recursive model and Discrete Fourier Transform are used in estimation of aggregate loss probabilities.

Phase type distributions, help us investigate the impact of frequency within frequency in estimation of aggregate losses. PH Poisson-Generalized Pareto model provided the best fit for Panjer class $(a,\ b,\ 0)$ while PH ZT Poisson-Generalized Pareto model provided the best fit for class $(a,\ b,\ 1)$ and PH two parameter Poisson Lindley-Generalized Pareto model provided the best fit for Poisson Lindley distributions. Finally, we propose phase type two parameter Poisson Lindley-Generalized Pareto as the best overall model for modeling secondary cancer data in Kenya and similar data. This research enables the insurance sector to improve its reserving models for cancer which has become a world wide menace.

Dedication

For Dad , Mum and Uncle Dr Nyenze
who supported and advised me all through.
You have been my greatest source of inspiration.
For Kim, Juli, Dan, Nicole, Dr Bulinda, Festo and Mary
for always believing in me and encouraging me to soldier on.

Acknowledgments

It is with pleasure to acknowledge the role of several individuals who were instrumental for completion of my PHD thesis.

First and foremost, I express my sincere gratitude to the Almighty God for giving me the knowledge, wisdom and strength to undertake and complete this research.

To Prof. Weke, I am entirely grateful for your support, guidance, encouragement and advice accorded during the writing of this thesis. Thank you for taking your time to read my work and for your brilliant comments. To Dr Bundi, thank you for your dedication and always being available to go through my work and guide me accordingly. Thank you for always encouraging me to do more and advising me on how I can make my work better. Thank you Prof. J.A.M Otieno, for your unwavering support, encouragement, wisdom, patience and enthusiasm throughout the whole process. I really thank you for your dedication to mold this research to its completion.

To the Faculty of Science and Technology and Department of Mathematics, I am appreciative of the support and guidance accorded to me by the academic staff during my progress seminar presentations and the support of non-academic staff. Special thanks to the Chairman Department of mathematics, Prof. Stephen Luketero for always being available when his support and guidance was required.

A special thanks to my colleagues in the department of mathematics and Kisii University as a whole for their moral support.

I cannot put in words how grateful I am to my mother and father for all the sacrifices they have made. Your prayers, love have been a pillar to hang on and have sustained me this far. To Kim, Juli, Dan, and Nicole thanks for always encouraging me and pushing me even to greater heights.

Cynthia Mwende Mwau

Nairobi, 2022.

Contents

Abbreviation and Notation

The abbreviations and notations used in this study are as listed below:

ZTP	Zero Truncated Poisson
ZTB	Zero Truncated Binomial
ZTG	Zero Truncated Geometric
MLE	Maximum Likelihood Function
DPH	Discrete Phase-type
CPHD	Compound Phase Type Distribution
PH PL	Phase Type Poisson Lindley
PH-P	Phase Type Poisson
PH-NB	Phase Type Negative Binomial
PH-B	Phase Type Binomial
PH-G	Phase Type Geometric
PH-OPPL	Phase type one parameter Poisson Lindley
PH-TPPL	Phase type two parameter Poisson Lindley
SBPD	Size-Biased Poisson Distribution
PH-ZTOPPL	Phase type zero truncated one parameter Poisson Lindley
PH-ZTTPPL	Phase type zero truncated two parameter Poisson Lindley
S_N	Compound distribution
X_i	Severity
N	Frequency
pdf	Probability density function
pgf	Probability generating
pmf	Probability Mass Function
cdf	Cumulative density function
cf	Characteristic function.
f_x	Severity probability distribution
p_n	Frequency probability distribution
P_n	Phase type distribution
DFT	Discrete Fourier Transform
FT	Fourier Transform
PH	Phase type
ETNB	Extended Truncated Negative Binomial
PLD	Poisson Lindley Distribution

List of Tables

List of Figures

1 INTRODUCTION

1.1 Background of the study

Aggregate loss distribution is distribution of aggregate monetary loss in activities which have occurred over a one year period. Different approaches have been developed to estimate aggregate losses, including Monte Carlo method, Panjer recursive formula, Fourier Transform and simulation [Pavel, 2011]. Aggregate losses are as:

$$S_N = \sum_{i=1}^{N} G_i \qquad S_N = G_1 + G_2 + \ldots + G_N \tag{1.1}$$

where:
G_i represents the claim severity/amount and N represents the claim count/number. Aggregate loss distributions were introduced way back in 1980's by [Dmitry et.al, 2001], [Glenn, 1981] , [Hewit, 1967], [Harry, 1981] and [Felix et. al, 2014]. Originally aggregate loss distributions were developed using ordinary distributions such as Poisson for claim frequency, Gamma and exponential distributions for claim severity distributions.

[Persi et. al, 1991] extended construction of aggregate loss distribution using mixtures such as Negative Binomial for claim frequency, Pareto and Generalized Pareto for claim severity distribution. [Mohamed et. al, 2010] further extended estimation of aggregate loss distribution using simulation method. Aggregate loss distributions have been evolving gradually with some of distributions being constructed only in their closed form hence they are calculated using numerical methods [Heckman et. al, 1983].

Modeling of aggregate losses plays an important role in decision making in the business field especially the insurance sector. The aspect of phase type distributions on aggregate loss distribution have not been widely investigated hence any development in this area provides great help to insurance firms.

1.1.1 Phase-type distribution

Phase-type distributions are derived by convoluting exponential distribution. These distributions come about from interrelated Poisson process which are in phases as shown in [Asmussen, 1996]. The sequence in which these processes occur is a stochastic process. It is shown by a random variable which describes the time until the Markov process reaches the one absorption state where the every state represents a phase.

Phase type distributions can be divided into two which are discrete phase-type distributions and continuous-time Markov process. Phase-type distributions which are discrete result from inter-related geometric distributions occurring in phases where the sequence the phases occur in can be stochastic processes. [Erlang, 1909] introduced phase type distributions and they have been extended to more modern theories by [Neuts, 1981] as well as [Asmussen, 1996]. Degenerate distribution (0 phase/empty phase), exponential distribution representing (1 phase), Erlang distribution representing (2 or more identical phases) were first introduced by [Erlang, 1909] and was later developed further by [Jensen, 1953]. Theoretical properties of phase type distributions have been studied by [O'cinneide, 1990] while [Asmussen, 1996] generalized risk models to model situations where the premium is depended on current reserves. He further extended his work to provide an algorithm to model finite time-horizon ruin probability. This was later extended to its application to survival analysis as well as queuing theory. The most recent work on phase type distribution has been done on its statistical inference where likelihood estimation is proposed based on the EM-algorithm and Markov chain Monte Carlo (MCMC) based approach. In the recent past phase-type distribution have been used to approximate any positive-valued distribution as they preserve the Markovian nature of the model which is crucial for tractable computation used in performance evaluation.

1.1.2 Markov models

A model of a random occurance which evolves from time to time in a manner which the past activities affects the future activities through the current activity with some degree of probability is known as a Markov chain. The "time" can be either be discrete that is integers, continuous that is real numbers or a set which is totally ordered. Markov chain describes a process which has been observed at discrete intervals. Markov model have been extended to modeled infectious diseases taking into consideration environmental factors which leads spreading of diseases which are infectious [Kehinde et. al, 2019]. This research tested the Markovian property and estimated how stationary the process was over the period.

Recent models have been developed to address spreading dynamics disease contagion and rumor spreading separately despite them being similar. This was elaborated by [Guilherme et. al, 2016] by developing a model based on discrete time Markov chain which included transitions which were plausible for both a disease contagion process and rumor propagation and consequently showed that their model covered traditional spreading schemes as well as features relevant in social dynamics, which include apathy, not remembering, and lost recovering of interest.

This was further advanced to joint observation and disease transition model which was modeled using latent continuous time Markov chain; and the observation process, according to a Markov-modulated Poisson process with observation rates that depend on the individual's underlying disease status by [Jane et. al, 2015]. All these types of matrices are considered in this research. [Sietske, 2009] showed that matrix-geometric distributions can equivalently be defined as distributions on the non-negative integers that have a rational probability generating function. It has been shown that the class of matrix-geometric distributions is strictly larger than the class of discrete phase-type distributions [Gareth et. al, 2015]. There are three special kinds of matrices which are: square matrix which has equal number of row, row vector which contains one row and colum vector which has one column as highlighted by [John et. al, 2016].

1.1.3 Panjer Recursive model

Panjer recursion are random variables of special type where in most general cases the distribution of S in equation (1.1) is a compound distribution. Recursion for special cases were introduced by Harry Panjer and they are used to calculate compound distribution. The claim size X_i is assumed to be i.i.d and independent of frequency distribution. For continuous severity distribution the severity probabilities are obtained by discretization of the claim density function but discrete severity distributions can be applied directly to Panjer recursive formula. The claim number N is a random variable which takes $0,\ 1,\ 2...$ values. Panjer recursion requires the probability of N to be a member of the Panjer class and the classes whose recursion has been developed are the $(a,\ b,\ 0)$ class and $(a, b, 1)$ class of distribution.

Distribution of class $(a, b, 0)$ should satisfy the relation:

$$P[N = k] = p_k \qquad p_k = \left(a + \frac{b}{k}\right) p_{k-1} \qquad k \geq 1 \tag{1.2}$$

where the initial value p_0 is determined such that $\sum_{k=0}^{\infty} p_k = 1$.
The distribution of class $(a, b, 1)$ should satisfy the relation;

$$P[N = k] = p_k \qquad p_k = \left(a + \frac{b}{k}\right) p_{k-1} \qquad k \geq 2 \tag{1.3}$$

where the initial value P_0 is an assumed value and p_1 is determined by:

$$p_1 + p_2 + ... + p_k = 1 - p_0 \tag{1.4}$$

1.1.4 Panjer class $(a, b, 0)$

Panjer class $(a,\ b,\ 0)$ recursion formula is satisfied by four distributins which are ,Binomial, Poisson, Geometric and Negative Binomial distributions. More distributions in this class can be derived by fixing the initial value p_j and consequently applying the recursion tp the subsequent vales. Negative Binomial distribution is constructed by fixing two parameters k and n using methods which rely on: Binomial expansion as well as Poisson with Gamma mixing distributions and mixing of iid random variables of Geometric distribution. Experiments with the random variable which represent the number of failures experienced before achiving the n^{th} success as well as the total number of trials which are needed to achieve the n^{th} success as highligted by [Oketch, 2011]. If the class $(a,\ b,\ 0)$ recursion holds for a given range of values of k, known distributions are then available. [Sundt et. al, 1981] proved that only these four distributions mentioned above belong to class $(a,\ b,\ 0)$. The four distributions can be represented by a united formula called united Panjer distribution .

The values of a, b and p_0 in these distributions can be represented depending on the distribution as:

i. Binomial distribution

$$a = \frac{-p}{1-p} \qquad b = \frac{p(n+1)}{1-p} \qquad p_0 = (1-p)^n$$

ii. Poisson distribution

$$a = \theta \qquad b = \lambda \qquad p_0 = e^{-\lambda}$$

iii. Negative Binomial distribution

$$a = 1-p \qquad b = (1-p)(r-1) \qquad p_0 = p^r$$

iv. Geometric distribution

$$a = 1-p \qquad b = 0 \qquad p_0 = p$$

The values of a, b, p_0 using united Panjer distribution of class $(a, b, 0)$ are expressed as:

$$a = \frac{\lambda}{\alpha+\lambda} \qquad b = \frac{(\alpha-1)\lambda}{\alpha+\lambda} \qquad p_0 = (1+\frac{\lambda}{\alpha})^{-\alpha}$$

A matrix has been derived from the recursion of Panjer class $(a, b, 0)$ for the distribution of compound distribution when frequency distribution belongs to the generalized Panjer class $(a, b, 0)$ family and this has been a major development in risk theory.

1.1.5 Panjer class (a,b,1)

Panjer recursion algorithm is used to estimate the probability distribution approximation of a compound random variable . Panjer class $(a, b, 1)$ contains distributions such as Zero truncated distributions as shown by [Fackler, 2009], Zero modified distribution as shown by [Younes, 2012], Extended truncated Negative binomial (ETNB) distribution and Sibuya distribution . Class $(a, b, 1)$ distributions increase the flexibility in modeling claim frequency distributions. The Zero modified distributions are derived from zero truncated distribution. In the case of zero truncated distribution as highlighted by [Elsayed, 2011] the value of zero is not recorded hence it can be expressed as:

$$p_k^T = \frac{1}{1-p_0} p_k \qquad k = 1, 2, 3, 4, \dots \tag{1.5}$$

Zero modified distribution is expressed as:

$$p_k^T = (1-p_0^m) p_k^T \qquad k = 1, 2, 3, 4, \dots \tag{1.6}$$

The Zero modified distributions can be derived from the distributions of class $(a, b, 0)$ by modifying formula (1.6) to:

$$p_k^m = \frac{1 - p_0^m}{1 - p_0} p_k \qquad k = 1, 2, 3, 4, \ldots \tag{1.7}$$

Panjer recursive formula of class $(a,\ b,\ 1)$ was derived by [Sundt et. al, 1981]. A matrix has been derived from the recursion of Panjer class $(a,\ b,\ 1)$ for the distribution of compound distribution when frequency distribution belongs to the generalized Panjer class $(a,\ b,\ 1)$ family. The development of a matrix form formula for the moment of compound distribution has also been a major step in evaluation of compound distribution.

1.1.6 Severity distributions

Severity distributions are distributions used to model claim count distributions. The Wei-bull distribution has been reviewed intesively since its was introduced in 1951 by Professor [Ernst, 1951]. [Mohammad, 2000] compared multiple methods of calculation of Weibull parameters which included how it fits based on the method of mean square error (MSE) and also the Kolmogorov- Smirnov (KS) criteria. Weibull distribution has been used in modeling squared returns of stock prices of the Cornerstone Insurance PLC and results showed that it has a good fit for the data . [Oscar, 1981] simplified method of moments in order to find Wei-bull distribution with specified mean and variance as it is tedious to calculate these values.

Discrete distributions are very important distriutions when modeling frequency data in various applied fields such as epidemiology, public health e.t.c. Parameter estimation of distributions can be considered in guiding on how to estimate parameters for the model which is very important for reliability engineers and applied statisticians. Ordinary discrete distributions such geometric and Poisson exhibit weak applicability in modeling failure times and frequency. This is majorly because most real frequency data will show either under-dispersion or over-dispersion. This is not the case with these distributions as highlighted by [Abdulhakim et. al, 2021]. [Rama et. al, 2015] worked on one parameter Poisson Lindley distribution in modelling frequency data and it arises from Poisson distribution when its parameter λ follows Lindley distribution. General expression for the r^{th} factorial moment of PLD has been obtained and hence its first four moments about origin has also been obtained by Shanker. One parameter Poisson Lindley distribution has been applied in data-sets relating to ecology and genetics to test its goodness of fit and the fit shows that it can be an important tool for modeling biological science data. [Rama et. al, 2016] obtained two-parameter Lindley distribution in 2013. When $\alpha = 1$ for two parameter Poisson Lindley then it becomes one parameter Poisson Lindley distribution. Two parameter Poisson distribution has been found to be a beNer model than the one parameter Poisson Lindley distribution for analyzing waiting time, survival time and grouped mortality data [Tanka, 2016]. Two parameter Poisson Lindley distribution have been estimated by [Rama et. al, 2016]. Two parameter Poisson Lindley distribution has been extended to three parameter Poisson. Parameters estimation of three state Poisson Lindley distribution has been explored using maximum like likelihood and method of moments and simulation study has been carried out to check the consistency of the maximum likelihood estimates. [Kishore et. al, 2018] applied three parameter Poisson distribution to read data and it was discovered that it is a flexible model that may be a useful alternative to known distributions like Poisson, Poisson Lindley, Two-parameter Poisson Lindley and many others for count data analysis . [Rama et. al, 2017] highlighted on Zero-truncated two parameter Poisson Lindley distribution.

1.2 Statement of the problem

Aggregate losses combine the likelihood (frequency) and size (severity) of losses. This produces a beNer estimate of the impact of the losses to the institution. Aggregate loss is usually expressed in terms of probability distribution to enable evaluation of the magnitude of the risk. [Harry, 1981] derived a recursive formula to estimate aggregate loss which required discrete claim frequency distributions and if claim amount distribution were continuous the values had to be discretized using method of rounding or method of local moment matching. [Xueyuan, 2010] introduced a different approach using matrix-form recursion which estimated compound distributions if claim severity ,X_i, were discrete or continuous phase-type distributions and claim frequency , N, were generalized class $(a, b, 0)$ family. The class of discrete phase-type distributions is one of the classes of distributions which are dense in the class of all discrete distributions how ever only few distributions have been explored.

Claim severity has been well explored in previous research while less has been done on claim frequency more so using phase type distributions which is a developing field. In this research we contribute in calculation of aggregate loss probabilities by in-cooperating phase type distributions in modeling claim frequency and consequently compound phase type distributions. The phase type distributions considered in this research will be constructed as well as their properties and parameter estimation methods developed. Phase type distributions requires one of its parameters to be a matrix hence a model to determine how to select the matrices will be developed and the proposed models applied to secondary cancer cases in Kenya to demonstrate their applicability and advantage. These models improve estimation of aggregate losses for cancer insurance policies as it in-cooperates transition of the secondary cancers. Phase type models are preferred because matrix parameters provide great flexibility. Cancer has over the years become one of the leading killer diseases hence insurance sectors have recently ventured into insurance policies to covers cancer patients. It is a dynamic disease hence it needs models that can capture its dynamic nature. In-cooperation of phase type distributions enables modeling of the dynamic aspect of cancer.

This research is aimed at developing a model which improves the estimation of aggregate losses of cancer and other diseases with transition states hence enabling insurance sectors to draft competitive policies which increase the uptake of cancer policies and other chronic diseases.

1.3 General objective

The primary research objective is to compute phase type compound probability generating function for various cases and estimate aggregate losses of secondary cancer cases.

1.3.1 Specific objectives

i. To develop phase type distributions using phase type Panjer class $(a, b, 0)$ and class $(a, b, 1)$ and calculate their properties.

ii. To formulate phase type Poisson Lindley distributions and determine their properties.

iii. To construct compound phase type probability generating functions when N is a phase type mixture and discrete phase type distribution.

iv. To develop a model of determining the matrix of the phase type distributions using multi-state Markov model.

v. To estimate aggregate loss probabilities of secondary cancer cases using the proposed phase type models.

1.4 Significance of the study

Distributions of aggregate claims are used to calculate premiums and estimate claim fluctuation reserves. The distribution of retained losses is useful for the insured in deciding on the degree of coinsurance expressed in deductible arrangements, stop loss levels and quota insurance. Phase type distributions provides great flexibility in estimating aggregate losses. In-cooperating phase type distributions helps insurance sector to almost accurately reserve for their anticipated losses hence reducing the risk of going into ruin. This enables the insurance companies to plan on different investment ventures which can increase returns hence increasing their capacity and creating more opportunities for the ambitious young generation.

Construction of phase type models enables great flexibility in estimation of aggregate losses be-cause of the matrix parameters. In this work phase type distributions are used as claim frequency distribution hence increasing the flexibility in estimation of claim frequency. Determination of an algorithm which can propose appropriate condition upon how to select proper matrices to build up claim number distributions which is developed in this research is a major development in estimation of aggregate loss distributions.

This achievement transforms the insurance sector in estimation of aggregate losses. Improving modeling of cancer policies transforms lives of policy holders who eventually suffer from cancer and help insurance sectors make beNer estimates in their computation. This enables more insurance firms introduce chronic illness insurance policies which eventually improves access to quality health care for cancer patients. Transforming the insurance sector gives me a sense of accomplishment as a mathematician as this improves the survival rate of cancer patients as it has been proven that cancer can be treated successfully if detected early and given the correct medication.

1.5 Data

We considered 850 patients between calender years 2013-2018 from a large health facility in Kenya. The following data about the patients was available, age, sex, type of cancer, date of diagnosis and transition stages of the cancer. This data-set was used to compute aggregate loss probabilities of secondary cancer cases and is shown in Appendices .

1.6 Organization of thesis

This thesis progresses as follows;chapter one give background of aggregate loss distributions as well as phase type distribution which chapter two highlights on previous literature done on aggregate losses and phase type distributions. Chapter three reviews distributions of class $(a, b, 0)$ and class $(a, b, 1)$ and also Panjer recursive model for these two classes. The pgf of these distributions which satisfy condition of class $(a, b, 0)$ and $(a, b, 1)$ are derived using pgf technique and iteration technique. Chapter four expresses distributions of class $(a, b, 0)$ as phase type distributions and also estimates their properties. Panjer recursive model for class $(a, b, 0)$ is expressed as a phase type recursive model. It also extends to distributions of class $(a, b, 1)$ and they are expressed as phase type distributions and also their properties are estimated. Panjer recursive model for class $(a, b, 1)$ are also expressed as phase type recursive model. These phase type distributions are used to model claim frequency for secondary cancer cases. Phase type Poisson mixture distributions and their properties are derived in chapter five. Chapter six derives expressions of the multi-state models using Chapman-Kolmogorov for three state, four state, five state and six state models for cancer data. These multi-state models represents the matrices in the phase type distributions used in modeling claim frequency.

Chapter seven estimates parameters of distributions used to model claim severity. Severity distributions considered are both discrete and continuous distributions. Parameters of severity distributions are estimated using method of moments. The continuous distributions are discretized in order to be applied in phase type Panjer recursive model for both class $(a, b, 0)$ and class $(a, b, 1)$ and in Discrete Fourier Transform. Severity distribution are used to model claim amounts of cancer cases and hence they are in-cooperated in estimating the aggregate losses. In chapter eight, transition probabilities and transition intensities of three state, four state, five state and six state are estimated. The long run probabilities which are represented as γ in our research are calculated for each multi-state model. Claim frequency probability for distributions of both classes are calculated as well as their moments. Claim severity probabilities are also estimated for each claim severity distribution. Aggregate loss probabilities in-cooperating different claim frequency distributions and severity distributions are estimated for class $(a, b, 0)$, class $(a, b, 1)$ and Poisson mixture distributions. Chapter nine concludes on our research and also highlight our recommendations.

2 LITERATURE REVIEW

2.1 Introduction

This chapter outlines relevant literature on aggregate losses as well as phase type models. This chapter is divided into subsections that explain on choice of frequency and severity distributions, compound distributions explored in previous research and outline the research gaps handled in this research.

2.2 Review framework

The literature is outlined as follows; aggregate losses distribution and methods applicable in calculation of aggregate losses are explored as well as framework used in choice of severity and frequency distributions. Previous literature on distributions used to model claim frequency are extensively explored as well as for severity distribution and previous research on aggregate loss distributions also examined. The distributions can either be ordinary distributions, mixture distributions or phase type distributions. Compound distributions explored so far are highlighted and the methods employed in calculation of the aggregate loss probabilities. This research is aimed at modeling claim data with multiple transition states hence phase type distributions are preferred to other models. Lastly areas of application of the phase type models are explored and their parameter estimation consequently identifying more research gaps.

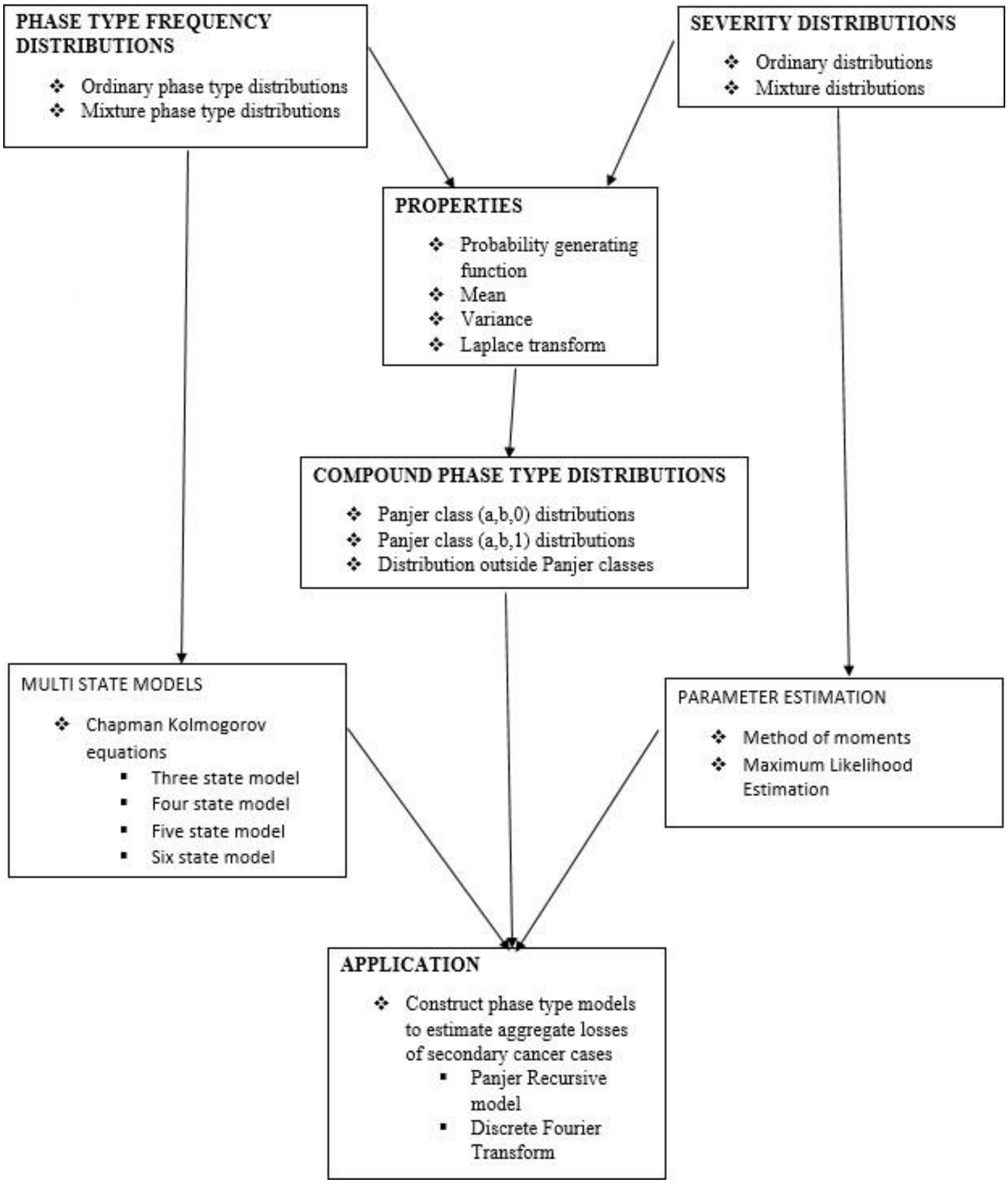


Figure 2.1. Model Framework

2.3 Aggregate losses

Aggregate loss distributions are calculated using compound distribution of claim frequency and claim severity distributions. Recursive formula for compound distribution (S) when claim frequencies, p_n ,belongs to the Schroters family was derived by Schroter (1991). The concept of compound distribution was further elaborated by [Harry, 1981] using a recursive formula when claim frequency (N) belong to the $(a, b, 0)$ family. This formula was generalized further to class $(a, b, 1)$ by [Sundt et. al, 1981]. [Eisele, 2006] derived a procedure for compound phase-type distribution but it was not computationally efficient because of its high convolutions. A simplified recursion algorithm for aggregate claims distribution with individual claim amounts having phase-type random variable has been derived by [Christian, 2006]. [Xueyuan, 2010] and [Kok et. al, 2010] derived phase type Panjer recursive model for class $(a, b, 0)$ and class $(a, b, 1)$, however they did not determine applicability of the models to real data and they did not develop a model to determine the matrix parameter. [Hess et. al, 2002] extended Panjer classes to class (a, b, k).

The claim frequency distribution which belong to class $(a, b, 0)$ are Poisson , Negative Binomial, Binomial and Geometric distributions [Garrido, 2006]. A frequency distribution that is a member of class $(a, b, 0)$ the following recursive relation must hold for some constant a and b.

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k} \qquad k = 1, 2, 3, \ldots \tag{2.1}$$

where p_k is the claim frequency probability. The initial probability p_0 of a member of class $(a, b, 0)$ is fixed implying that the sum of all p_k must sum to 1. The member of class $(a, b, 0)$ class has two parameters a and b [Stuart et. al, 2008].

[Kok et. al, 2010] extended [Xueyuan, 2010] work expressed a matrix-form recursion formula for class $(a, b, 1)$ family. The mean and variance of these distributions were evaluated. [Kok et. al, 2010] also discussed the mixtures of zero modified and zero truncated versions of logarithmic distributions and their linear combinations. It is clearly demonstrated in this work how to compute moments of compound distribution (S) recursive based on aggregate claim distribution. Only some very special members of the generalized $(a, b, 1)$ family are examined in this work which opens more problems to be explored. Panjer class $(a, b, 1)$ contains distributions such as Zero truncated distributions ($p_0 = 0$), Zero modified distribution ($p_0 > 0$) , Extended truncated Negative binomial(ETNB) distribution and Sibuya distribution .

Class $(a, b, 1)$ distributions increase the flexibility in modeling claim frequency distributions [Marcelo et. al, 2015]. The Zero modified distributions are derived from zero truncated distribution where in the case of zero truncated distribution the value of zero is not recorded. Distributions which satisfy conditions of Panjer class $(a, b, 1)$ include:

(a). Zero-truncated distribution,where $P_0 = 0$

i. Zero truncated Poisson distribution.

ii. Zero truncated Binomial distribution.
iii. Zero truncated Geometric distribution.

(b). Zero-modified distributions where $P_0 = 0$

i. Zero modified Poisson distribution.
ii. Zero modified Binomial distribution.
iii. Zero modified Geometric distribution.
iv. Zero modified Extended truncated negative binomial distribution(ETNB).

(c). Other distributions

i. Extended truncated Negative binomial (ETNB) distribution.
ii. Logarithmic distribution.
iii. Sibuya distribution.

Panjer recursive formula of class $(a, b, 1)$ was derived by[Sundt et. al, 1981]. A matrix has been derived from the recursion of Panjer class $(a, b, 1)$ for the distribution of compound distribution when frequency distribution belongs to the generalized Panjer class $(a, b, 1)$ family. Generalized phase type distributions have been derived for class $(a, b, 0)$ distributions . [Vilar et. al, 2009] assumed that insurance companies provides historical sample of claim frequency and claim severity hence he developed a non-parametric approach which can be used in insurance. [Harchol, 2012] based his work on non parametric estimators to calculate density functions where data is censored or truncated using Monte Carlo simulation methods and bootstrap re-sampling. A methodology useful in comparing various strategies used in pricing insurance products was developed. Different numerical methods which can be used in calculation of aggregate loss so far are:

(i) Monte Carlo simulations.
(ii) Panjer recursion.
(iii) Heckman-meyers.
(iv) Fourier Transform.

Panjer recursive model is the oldest method of calculating aggregate loss distribution and is discussed in [Heckman et. al, 1983]. The method of Heckman-Meyers is discussed in detail in Heckman [Heckman et. al, 1983] and its application in calculation of aggregate loss models. The application of Fast Fourier Transform in calculation of aggregate loss models is discussed in [Robertson, 1992]. The latest method to be developed is the stochastic simulation by [Mohamed et. al, 2010]. [Pavel, 2010] reviewed the three numerical algorithms which had been developed by 2010. [Paul et.al, 2010] showed that Monte Carlo method is the easiest to implement but it is a slow method hence Panjer recursive method and Fourier Transforms are more preferred. [Kumer et. al, 2011] reviewed and extended Panjer's recursion formula used in derivation of compound negative binomial distributions where they explored gamma as well as its mixtures distributions and developed a theory which can apply R software directly. The accuracy of the method used proved to be beNer and the computation time quite faster. The numerical methods considered in this research are Panjer recursive model as some of the distribution considered follow Panjer classes and Discrete Fourier Transform for distributions that do not follow Panjer classes.

A model to determine matrix parameters is developed in this research and extends application of Panjer class $(a, b, 1)$ distributions not considered so far as well as mixture distributions. [Sundberg, 1974] highlighted that aggregate loss models can be modeled using two modeling approaches which are:

(i) **Individual risk model**
This model emphasizes the loss from each individual contract and represents the aggregate losses as:

$$S_n = G_1 + G_2 + ... + G_n$$

where $G_i(i = 1, 2, 3, ...n)$ is the loss amount, n denotes fixed number of contracts in the portfolio which is a fixed number and G_i are independent and not necessarily identically distributed.

(ii) **Collective Risk model**
This model represents aggregate losses in terms of a frequency distribution and a severity distribution and is expressed as:

$$S_N = G_1 + G_2 + ... + G_N$$

where N is a random number representing the number of losses or payments, $G_i(i = 1, 2, 3, ..., N)$ represents the claims amounts which are assumed to be iid and both N and G_i are assumed to be independent of each other .

Severity and frequency distribution can be applied to the collective risk model to determine the aggregate losses. Severity distributions have been well explored in previous literature, however less research has been done on impact of claim frequency on aggregate losses. This research expands the scope of claim frequency distributions in estimation of aggregate losses. Phase type distributions are employed to address this short coming.

2.4 Frequency distribution

Frequency distributions can be divided into three categories which are discrete distributions, continuous distributions and mixed distributions.

2.4.1 Discrete distributions

A discrete probability distribution depicts the occurrence of discrete outcomes and is made up of discrete variables. The probability mass function of discrete distributions can be expresses as:

$$p(x) = \sum_{x=0}^{n} f(x)$$

Most claim count distributions considered so far are discrete distributions. This research extend used of discrete distributions as claim count distributions to phase type discrete count distributions in estimating aggregate losses.

2.4.2 Continuous distributions distribution

A continuous probability distribution is a distribution in which the random variable can take on any continuous value . The probability density function is expressed as:

$$p(x) = \int_{x=0}^{n} f(x)\,dx$$

2.4.3 Mixed distributions

A mixture distribution is a probability distribution whose random variable is derived from a group of other random variables. The random variable is selected randomly from the collection of other random variables according to the given probabilities of selection. The values of selected random variables are then realized. The underlying random variables can be random real numbers and if they are continuous then the outcome will be continuous. If the underlying random variable is discrete then the outcome is a discrete. Discrete mixture are expresses as:

$$p(x) = \sum_{x=0}^{n} w_i f(x)$$

Continuous mixture are expresses as:

$$p(x) = \int_{x=0}^{n} w_i f(x)\,dx$$

This research extend used of mixture distribution in modeling claim count data to phase type mixture distributions.

2.4.4 Phase type distributions

Phase type distributions are constructed by mixing of exponential distributions resulting from a chain of inter-related Poisson processes which occur in sequence also known as phases. [Sophie et. al, 2012] showed that under the constraint that their representation is to be non-negative, Poisson distributions which are extensions of phase type disributions. [Gabor et. al, 2012] introduced the libphprng library for generating random-variates from PH distributions. A compound random variable X is important in various experiences which include reliability which is used for modeling the lifetime of a system in a particular shock model as shown in [Verbelen, 2013].

Research has been done on distribution of the random variable $T = \sum_{i=1}^{N} G_i$ where G_i for $i \geq 1$ is a sequence made random variables which are independent and identically distributed and also have a common phase-type distribution with the distribution of T having been obtained using phase-type distributions as shown in [Serkan, 2016]. Victoria generalized phase type distribution by replacing the underlying by Markov mixtures process enabling them to model heterogeneity and inclusion of past information which is due to the Markov property of the underlying process as shown in [Surya, 2016].

[Asger et. al, 2018] demonstrated how similar coalescent theory and Phase-type theory are and the implication of that close relation. The translations are useful, complex and difficult to derive. Coalescent theory formula equations are trivial to define and compute using phase-type theory as well as the matrix notation. This work obtained explicit formula for the joint distribution of the height of a tree, explicit formula for expected values, co-variances of the height of a tree height and total length of branchs. [O'cinneide, 1990] highlighted on Phase type distributions and their invariant polytopes. [Mogens, 2005] introduced use of phase type distributions in risk theory while [Antony et. al, 2020] developed compound distribution to model extreme natural disasters in Kenya. Continuous phase-type distribution have various special cases which are:

(i) **Degenerate distribution**- this is where the point mass is zero or it is an empty phase-type distribution meaning it has 0 phases.

(ii) **Erlang distribution** – This has 2 or more phases which are identical and in sequence.

(iii) **Deterministic distribution (or constant)**– This is limiting case of the Erlang distribution where the number of phases consequently become infinite and the time in each and every state becomes zero.

(iv) **Coxian distribution** – This has 2 or more phases which are not necessarily identical in their sequence with a certain probability of transiting to the absorbing state after each phase.

(v) **Hypoexponential distribution** – it has two or more phases which are in sequence and they can either be non-identical or a mixture of identical and non-identical phases. [Mogens et. al, 2017].

This research modifies the concept of Hypoexponential distribution which is expressed as a phase type distribution as:

$$\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \ldots, \alpha_n)$$

with $\sum_{i=1}^{n} \alpha_i = 1$ and

$$S = \begin{bmatrix} -\lambda_1 & 0 & 0 & 0 & 0 \\ 0 & -\lambda_2 & 0 & 0 & 0 \\ 0 & 0 & -\lambda_3 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_4 & 0 \\ 0 & 0 & 0 & 0 & -\lambda_5 \end{bmatrix}.$$

2.5 Multi state models

Phase-type distribution is a probability distribution which is constructed by mixing exponential distributions. Markov processes are stochastic process and they are shown by a random variable which describes the time up to absorption for a Markov process which has one absorbing state where each state represents one of the phases.

A mathematical model representing of a random scenario which evolves with time in a manner that the past affects the future through the current state with some level of probability. Markov chain describes a process observed at discrete intervals and they have been used in epidemiological models as shown by [Abdulhakim et. al, 2021]. However they do not project other relevant disease information such as probability of geNing infection and recovery from first infection or the anticipated time to geNing an infection and recovering for both prone and infected people [Juan, 2020]. These disease models considered in estimating the transition probabilities (TPs) cannot generalize the transition estimates of disease outcomes at discrete time steps for predictions in the future. This necessitated the adoption of a discrete-time Markov chain model. [Clement et. al, 2019] sought to address aforementioned issues using discrete-time Markov chain model. This study came to the conclusion that hepatitis B is more infectious over a long period of time than tuberculosis or HIV despite the probability of geNing first infection of these two diseases being comparatively low within the population considered. Hiv infected Patients had a considerably lower life expectancy compared to those suffering from tuberculosis and Hepatitis B. This study came to the conclusion that discrete-time Markov models are good models in modeling diseases dynamics in Ghana .

[Asmussen, 1996] developed a model based on discrete time Markov chain including all transitions affecting disease contagion process as well as rumor propagation. They consequently showed that their model covered traditional spreading schemes as well as features which are relevant in social dynamics, such as apathy, in ability to remember, and lost recovering of interest. [Manuel et. al, 2021] considered a non-homogeneous continuous time Markov chain model for Long-Term Care to monitor the quality of the labeling using Portuguese life expectancies taking into consideration reasonable monthly costs for each dependence state and consequently computing them by Monte Carlo simulation, trajectories of the Markov chain process hence deriving relevant information for model validation and premium calculation. Chapman-kolmogorov equation is used to derive the Multi state models considered in this research. It based on the fact that, a process which begins in state i at time s and is in state j at time r occurs through some state $k \in M$ at an unknown intermediate time R i.e

$$p_{ij}(s,r) = \sum_{k=1}^{n} pr[Z(s,r) = j, Z(s,r) = k|Z(s,s) = i] \qquad p_{ij}(s,r) = \sum_{k=1}^{n} p_{ik}(s,R)p_{kj}(R,r) \qquad \textbf{(2.2)}$$

In the recent past insurance companies have embraced the thought of insurance for chronic diseases such as cancer hence multi state models improves it's estimation as it allows modeling of transition between states.

3 PHASE TYPE DISTRIBUTIONS OF CLASS (a,b,0) AND CLASS (a,b,1)

3.1 Introduction

Discrete phase type distributions (DPHD) are derived from a Markov chain which has one which is absorbing and retains the remaining states as transient. The transition matrix of a phase type distribution is expressed as:

$$C = \left(\begin{array}{cccccc|c} c_{10} & c_{11} & c_{12} & c_{13} & \cdots & c_{1m} \\ c_{20} & c_{21} & c_{22} & c_{23} & \cdots & c_{2m} \\ c_{30} & c_{31} & c_{32} & c_{33} & \cdots & c_{3m} \\ \vdots & \vdots & \vdots & \ddots & . & . \\ c_{m0} & c_{m1} & c_{m2} & c_{m3} & \cdots & c_{mm} \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$= \begin{pmatrix} Z & \underline{y} \\ \underline{0} & 1 \end{pmatrix}$$

where

y represents $(n * 1)$ non-zero column vector.

Z represents $(n * n)$ non-zero matrix.

0 represents $(1 * n)$ zero row vector.

1 represents $(1 * 1)$ single element matrix.

The higher orders of C are represented as:

$$C^2 = CC = \begin{pmatrix} Z & \underline{y} \\ \underline{0} & 1 \end{pmatrix} \begin{pmatrix} Z & \underline{y} \\ \underline{0} & 1 \end{pmatrix}$$

$$= \begin{pmatrix} Z^2 & B_2 \\ \underline{0} & 1 \end{pmatrix}$$

Hence C^n can be expressed as:

$$C^n = \begin{pmatrix} Z^n & B_n \\ \underline{0} & 1 \end{pmatrix} \quad \text{where } B_n = \left(\sum_{k=0}^{n-1} Z^k \right) \underline{u}$$

In this chapter we improve on the work of [Xueyuan, 2010]. The distributions considered in this chapter and chapter four satisfy the conditions of discrete phase type transition matrix. The algorithm of selecting matrix parameters for phase type distributions was not determined which will be one of the main objectives of this study. In this chapter distributions of class $(a, b, 0)$ and class $(a, b, 1)$ are expressed as phase type distributions and their properties derived.

3.2 PH Panjer class (a,b,0) distribution and its aggregate loss distribution recursively

3.2.1 PH Panjer class $(a, b, 0)$ distribution using pgf technique

PH Panjer formula for class $(a, b, 0)$ is illustrated as:

$$P_n = \left(A + \frac{B}{n} \right) P_{n-1} \tag{3.1}$$

Multiplying all through by n results to:

$$\begin{aligned} nP_n &= (An + B) P_{n-1} \\ &= \left((n-1+1)A + B \right) P_{n-1} \\ &= (n-1) AP_{n-1} + (A+B) P_{n-1} \qquad n = 1, 2, \ldots \end{aligned} \tag{3.2}$$

Expressing equation (3.2) in terms of pgf and multiplying it by s^n and sum the result over n.

$$\sum_{n=1}^{\infty} nP_n s^n = A \sum_{n=1}^{\infty} (n-1) P_{n-1} s^n + (A+B) \sum_{n=1}^{\infty} P_{n-1} s^n \tag{3.3}$$

Factoring out s, s^2 and s respectively results to:

$$s \sum_{n=1}^{\infty} nP_n s^{n-1} = As^2 \sum_{n=2}^{\infty} (n-1) P_{n-1} s^{n-2} + (A+B) s \sum_{n=1}^{\infty} P_{n-1} s^{n-1} \tag{3.4}$$

Divide through by s resulting to:

$$\sum_{n=1}^{\infty} nP_n s^{n-1} = As \sum_{n=2}^{\infty} (n-1) P_{n-1} s^{n-2} + (A+B) \sum_{n=1}^{\infty} P_{n-1} s^{n-1} \tag{3.5}$$

Let the pgf of phase type frequency distribution be defined as:

$$Y(s) = \sum_{n=0}^{\infty} P_n s^n$$

The first derivative of the pgf is:

$$Y'(s) = \sum_{n=1}^{\infty} nP_n s^{n-1} \tag{3.6}$$

Let the pgf of the phase type frequency distribution be defined as:

$$Y(s) = \sum_{n=1}^{\infty} P_{n-1} s^{n-1} \tag{3.7}$$

The first derivative of the pgf is:

$$Y'(s) = \sum_{n=2}^{\infty} (n-1)P_n s^{n-2} \tag{3.8}$$

Combining equation (3.6),equation (3.8) and equation (3.5) it becomes :

$$\begin{aligned} Y'(s) &= AsY'(s) + (A+B)Y(s) \\ (I - As)Y'(s) &= (A+B)Y(s) \end{aligned} \tag{3.9}$$

Rearranging equation (3.9) it becomes:

$$\frac{Y'(s)}{Y(s)} = \quad A+B \quad I-As \quad ^{-1} \tag{3.10}$$

This expresses the discrete distributions of phase type Panjer of class $(a, b, 0)$ in terms of pgf.

Theorem 3.2.1 (PH Panjer class $(a, b, 0)$ distributions). *The distributions arising from pgf of phase type Panjer class $(a, b, 0)$ model*

$$\frac{Y'(s)}{Y(s)} = \quad A+B \quad I-As \quad ^{-1} \tag{3.11}$$

are:

(i) *Phase type Poisson when* $A = 0, B > 0$.

(ii) *Phase type Negative binomial* $A > 0, B > 0$.

(iii) *Phase type Binomial when* $BA^{-1} > 0, BA^{-1} < 0$.

(iv) *Phase type Geometric when* $A > 0, B = 0$.

Proof of theorem 3.2.1

(i) **When** $A = 0, B = 0$

From equation (3.11) replacing the value of A and the value of B results to:

$$\frac{Y^1(s)}{Y(s)} = B$$

$$\int d\, lnY(s) = \int B\, ds$$

$$lnY(s) = Bs + C$$

The above equation simplifies to:

$$Y(s) = e^C e^{Bs} \quad \textbf{(3.12)}$$

Replacing s with 1 results to :

$$Y(1) = e^C e^B \qquad 1 = e^C e^B \qquad e^C = e^{-B} \quad \textbf{(3.13)}$$

Combining equation (3.13) and equation (3.12) results to:

$$Y(s) = e^{-B} e^{Bs} \qquad Y(s) = e^{-B+Bs} \qquad Y(s) = e^{B(s-1)} \quad \textbf{(3.14)}$$

Summary

This represents the pgf of phase type Poisson distribution with parameter B.

Theorem 3.2.2 (**Properties of phase type Poisson distribution**). *Letting B from equation (3.14) be* Λ *the properties of Phase type Poisson can be expressed as:*

(a) pgf

$$Y(\hat{s}) = \Upsilon e^{\Lambda(s-1)} \mathbf{1}^T \quad \textit{(3.15)}$$

(b) Expectation

$$E(\hat{N}) = \Upsilon \Lambda \mathbf{1}^T \quad \textit{(3.16)}$$

(c) Variance

$$Var(\hat{N}) = \Upsilon \Lambda \mathbf{1}^T \quad \textit{(3.17)}$$

Proof of theorem 3.2.2

(a) The pmf of phase type Poisson distribution is:

$$y(\hat{n}) = \frac{e^{-\Lambda}\Lambda^n}{n!}$$

The pgf of Poisson is given by:

$$Y(s) = \sum_{n=0}^{\infty} y(\hat{n}) s^n \qquad = \sum_{n=0}^{\infty} \frac{e^{-\Lambda}\Lambda^n}{n!} s^n \qquad = e^{-\Lambda} \sum_{n=0}^{\infty} \frac{(s\Lambda)^n}{n!}$$

$$Y(s) = e^{-\Lambda} e^{\Lambda s} \qquad Y(s) = e^{\Lambda(s-1)} \tag{3.18}$$

Equation (3.18) can be transformed to a proper pgf by multiplying by Υ on the left hand side and 1^T on the right hand side to become:

$$Y(\hat{s}) = \Upsilon e^{\Lambda(s-1)} 1^T \tag{3.19}$$

Derivative of pgf can be used to obtain other properties of PH Poisson distribution. The derivatives of equation (3.18) are obtained as:

$$Y'(s) = \Lambda e^{\Lambda(s-1)} \tag{3.20}$$

$$Y''(s) = \Lambda^2 e^{\Lambda(s-1)} \tag{3.21}$$

In general the k^{th} factorial moments is given by:

$$Y^k(s) = \Lambda^k e^{\Lambda(s-1)}$$

(b) $E(N) = Y'(1)$ hence equation (3.20) can be expressed as:

$$Y'(s) = \Lambda e^{\Lambda(s-1)} \qquad Y'(1) = \Lambda \qquad E(N) = \Lambda \tag{3.22}$$

Equation (3.22) can be wriNen as a proper expectation as:

$$E(\hat{N}) = \Upsilon \Lambda 1^T \tag{3.23}$$

(c) Variance of N can be expressed as:

$$Var(N) = Y''(1) + Y'(1) - [Y'(1)]^2 \tag{3.24}$$

$Y''(1)$ can be obtained from equation (3.21) as:

$$Y''(s) = \Lambda^2 e^{\Lambda(1-1)} \qquad Y''(s) = \Lambda^2 \tag{3.25}$$

Hence equation (3.24) can be expressed as:

$$Var(N) = \Lambda^2 + \Lambda - [\Lambda]^2 \qquad Var(N) = \Lambda \tag{3.26}$$

Equation (3.26) can be rewriNen as :

$$Var(\hat{N}) = \Upsilon \Lambda 1^T \tag{3.27}$$

(ii) **When** $A > 0, B > 0$

From equation (3.10) replacing A and B results to :

$$\frac{Y'(s)}{Y(s)} = (A+B)(I-As)^{-1}$$

$$\int d\ lnY(s) = \int (A+B)(I-As)^{-1}\ ds$$

Introduce $-A$

$$\int d\ lnY(s) = (A+B)(-A)^{-1} \int (I-As)^{-1}\ ds$$

$$lnY(s) = (A+B)(-A)^{-1} ln(I-As) + ln\,C \qquad lnY(s) = ln(I-As)^{-[(A+B)(A)^{-1}]} + ln\,C$$

$$lnY(s) = ln\,C\ (I-As)^{-[(A+B)(A)^{-1}]}$$

Therefore:

$$Y(s) = [C]\ (I-As)^{-[(A+B)(A)^{-1}]} \tag{3.28}$$

Let $s = 1$

$$Y(1) = C\ (I-A)^{-[(A+B)(A)^{-1}]} \qquad C = (I-A)^{(A+B)(A)^{-1}} \tag{3.29}$$

Replacing equation (3.29) in equation (3.28) :

$$Y(s) = \left[(I-A)(I-As)^{-1}\right]^{(A+B)(A)^{-1}} \tag{3.30}$$

Summary

This is the pgf of phase type negative binomial distribution with parameters $(A+B)(A)^{-1}$ and $(I-A)$ for $0 < A < I$. Any data set that has the values of $A > 0$ and $B > 0$, the most appropriate frequency distribution is the negative binomial distribution.

Theorem 3.2.3 (Properties of phase type Negative Binomial distribution). *Letting A from equation (3.30) be Q and* $B = (\alpha - 1)Q$ *the properties of phase type Negative Binomial distribution are expressed as:*

(a) pgf

$$Y(\hat{s}) = \Upsilon\{[I-Q][I-sQ]^{-1}\}^{\alpha}\mathbf{1}^T \tag{3.31}$$

(b) Expectation

$$E(\hat{N}) = \alpha\Upsilon Q[I-Q]^{-1}\mathbf{1}^T \tag{3.32}$$

(c) Variance

$$Var(\hat{N}) = \alpha\Upsilon\{Q^2[I-Q]^{-2} + Q[I-Q]^{-1}\}\mathbf{1}^T \tag{3.33}$$

Proof of theorem 3.2.3

(a) *The pmf of phase type Negative Binomial distribution is:*

$$y(\hat{n}) = \binom{n+\alpha-1}{n}[I-Q]^{\alpha}Q^{n} \qquad y(\hat{n}) = \binom{n+\alpha-1}{n}P^{\alpha}Q^{n}$$

The pgf of n is given by:

$$Y(s) = \sum_{n=0}^{\infty} y(n)s^{n} = \sum_{n=0}^{\infty}\binom{n+\alpha-1}{n}P^{\alpha}Q^{n}s^{n} = P^{\alpha}\sum_{n=0}^{\infty}(-1)^{n}\binom{-\alpha}{n}(Qs)^{n}$$

$$Y(s) = P^{\alpha}\left[I-Qs\right]^{-\alpha} = \left[P[I-Qs]^{-1}\right]^{\alpha} \tag{3.34}$$

P can be expressed in terms of Q as $P = [I - Q]$ *hence equation (3.34) can be written as:*

$$Y(s) = \left([I-Q][I-sQ]^{-1}\right)^{\alpha} \tag{3.35}$$

Equation (3.35) can be transformed to a proper pgf by multiplying by Υ *on the left hand side and* $\mathbf{1}^{T}$ *on the right hand side to become:*

$$Y(\hat{s}) = \Upsilon\left([I-Q][I-sQ]^{-1}\right)^{\alpha}\mathbf{1}^{T} \tag{3.36}$$

which is the pgf of Phase type Negative Binomial distribution. Derivative of pgf can be used to obtain other properties of phase type Negative Binomial distribution. The derivatives of equation (3.34) are obtained as:

$$Y'(s) = \alpha P^{\alpha}Q[I-sQ]^{-\alpha-1} \tag{3.37}$$

$$Y''(s) = \alpha(\alpha+1)P^{\alpha}Q^{2}[I-sQ]^{-\alpha-2} \tag{3.38}$$

In general the k^{th} *factorial moments is given by:*

$$Y^{k}(s) = \alpha(\alpha+1)(\alpha+2)\ldots(\alpha+k-1)P^{\alpha}Q^{k}[I-sQ]^{-\alpha-k}$$

(b) $E(N) = Y'(1)$ *hence equation (3.37) can be expressed as:*

$$Y'(s) = \alpha P^{\alpha}Q[I-Q]^{-\alpha-1} = \alpha P^{\alpha}Q[I-Q]^{-\alpha}[I-Q]^{-1}$$

Hence:

$$E(N) = \alpha Q[I-Q]^{-1} \tag{3.39}$$

Equation (3.39)can be written as a proper expectation as:

$$E(\hat{N}) = \alpha\Upsilon Q[I-Q]^{-1}\mathbf{1}^{T} \tag{3.40}$$

(c) Variance of N can be expressed as shown in (3.24) and $Y''(1)$ can be obtained from equation (3.38) as:

$$Y''(1) = \alpha(\alpha+1)P^{\alpha}Q^{2}[I-Q]^{\alpha-2} \qquad = \alpha(\alpha+1)[I-Q]^{\alpha}Q^{2}[I-Q]^{\alpha-2}$$

$$Y''(1) = \alpha(\alpha+1)Q^{2}[I-Q]^{-2} \tag{3.41}$$

Hence equation (3.24) can be expressed as:

$$Var(N) = \left(\alpha(\alpha+1)Q^{2}[I-Q]^{-2} + \alpha Q[I-Q]^{-1} - \alpha^{2}Q^{2}[I-Q]^{-2}\right)$$
$$= \alpha Q^{2}[I-Q]^{-2} + \alpha Q[I-Q]^{-1}$$

$$Var(N) = \alpha\left(Q^{2}[I-Q]^{-2} + Q[I-Q]^{-1}\right) \tag{3.42}$$

Equation (3.42) can be rewritten as :

$$Var(\hat{N}) = \alpha Y\left(Q^{2}[I-Q]^{-2} + Q[I-Q]^{-1}\right)1^{T} \tag{3.43}$$

(iii) **When** $A < 0, B > 0$

From equation (3.10) replacing the A and B results to:

$$\frac{Y'(s)}{Y(s)} = (A+B)(I - As)^{-1}$$

From equation (3.30) we know that:

$$Y(s) = \left[(I-A)(I-As)\right]^{(A+B)A^{-1}}$$

Let $(A+B)(A)^{-1} = -M$

Where M is a positive matrix .

Therefore:

$$Y(s) = \left[(I-A)(I-As)^{-1}\right]^{-m} \qquad = \left[(I-As)(I-A)^{-1}\right]^{m}$$

$$Y(s) = \left(I(I-A)^{-1} - [A(I-A)^{-1}]s\right)^{m} \qquad Y(s) = \left((I-A)^{-1} + \left(-A(I-A)^{-1}\right)s\right)^{m} \tag{3.44}$$

Summary

This is the pgf of phase type binomial distribution with matrix parameter M. This means that for any data with matrix of value $A < 0$ and matrix of value $B > 0$ the most appropriate frequency distribution for that data set is phase type binomial distribution.

Theorem 3.2.4 (**Properties of phase type Binomial distribution**). *Letting A from equation (3.44) be* $-(I-Q)Q^{-1}$ *and* $B = (\alpha+1)(I-Q)Q^{-1}$ *the properties of phase type Binomial distribution are expressed as:*

a) pgf

$$Y(\hat{s}) = \Upsilon[Ps + Q]^{\alpha}1^T \quad (3.45)$$

b) Expectation

$$E(\hat{N}) = \alpha\Upsilon[I - Q]1^T \quad (3.46)$$

c) Variance

$$Var(\hat{N}) = \alpha\Upsilon\{[I - Q]Q\}1^T \quad (3.47)$$

Proof of theorem 3.2.4

(a) The pmf of phase type Binomial distribution is:

$$y(\hat{n}) = \binom{\alpha}{n}(I - Q)^n Q^{\alpha-n} \qquad y(\hat{n}) = \binom{\alpha}{n} P^n Q^{\alpha-n}$$

The pgf of n is given by:

$$Y(s) = \sum_{n=0}^{\infty} y(\hat{n})s^n = \sum_{n=0}^{\infty}\binom{\alpha}{n} P^n Q^{\alpha-n} s^n = (Ps + Q)^{\alpha} \quad (3.48)$$

Equation (3.48) can be transformed to a proper pgf by multiplying by Υ on the LHS and 1^T on the RHS to become:

$$Y(\hat{s}) = \Upsilon[Ps + Q]^{\alpha}1^T \quad (3.49)$$

which is the pgf of Phase type Binomial distribution. Derivative of pgf can be used to obtain other properties of phase type Binomial distribution. The derivatives of equation (3.48) are obtained as:

$$Y'(s) = \alpha P[Ps + Q]^{\alpha-1} \quad (3.50)$$

$$Y''(s) = \alpha(\alpha - 1)P^2[Ps + Q]^{\alpha-2} \quad (3.51)$$

In general the k^{th} factorial moments is given by:

$$Y^k(s) = \alpha(\alpha - 1)(\alpha - 2)\ldots(\alpha + k - 1)P^k[Ps + Q]^{\alpha-k}$$

(b) $E(N) = Y'(1)$ hence equation (3.50) becomes:

$$Y'(1) = \alpha P[P + Q]^{\alpha-1} = \alpha P = \alpha[I - Q]$$

Hence:

$$E(N) = \alpha[I - Q] \quad (3.52)$$

Equation (3.52) can be wriNen as a proper expectation as:

$$E(\hat{N}) = \alpha\Upsilon[I - Q]1^T \quad (3.53)$$

(c) Variance of N can be expressed as shown in (3.24) and $Y''(1)$ can be obtained from equation (3.51) as:

$$Y''(1) = P^2\alpha(\alpha-1)[P+Q]^{\alpha-2} \qquad Y''(1) = P^2\alpha(\alpha-1) \qquad \textbf{(3.54)}$$

Hence equation (3.24) can be expressed as:

$$Var(N) = \{P^2\alpha(\alpha-1) + P\alpha - P^2\alpha^2\} \qquad = P\alpha[I-P] \qquad = P\alpha Q \qquad \textbf{(3.55)}$$

Equation (3.55) can be rewriNen as :

$$Var(\hat{N}) = \alpha\Upsilon PQ1^T \qquad \textbf{(3.56)}$$

(iv) **When** $A > 0, B = 0$

From equation (3.11) replacing A and B it becomes :

$$\frac{Y'(s)}{Y(s)} = (A)(I-As)^{-1} \qquad \frac{d}{ds}lnY(s) = (A)(I-As)^{-1} \qquad \int d\,lnY(s) = \int (A)(I-As)^{-1}ds$$

Introduce a negative sign:

$$\int dlnY(s) = -\int (-A)(I-As)^{-1}\,ds \qquad lnY(s) = -ln(I-As) + ln\,C$$

Therefore:

$$Y(s) = (C)(I-As)^{-1} \qquad \textbf{(3.57)}$$

Let $s = 1$

$$Y(1) = C(I-A)^{-1} \qquad I = C(I-A)^{-1} \qquad C = I-A \qquad \textbf{(3.58)}$$

Combining equation (3.58) and equation (3.57) results to:

$$Y(s) = (I-A)(I-As)^{-1} \qquad \textbf{(3.59)}$$

Summary

This is the pgf of a phase type geometric distribution with parameter $(I-A)$.

$$i.e \;\; P_n = A^n(I-A) \qquad n = 0, 1, 2, \ldots$$

For any data which has its value of $B = 0$, the suitable frequency distribution is phase type geometric distribution and its mean should be greater than its variance.

Theorem 3.2.5 (Properties of phase type Geometric distribution). *Letting A from equation (3.59) be Q and* $B = 0$ *the properties of phase type Geometric distribution are expressed as:*

(a) pgf

$$Y(\hat{s}) = \Upsilon[I - sQ]^{-1}[I - Q]\mathbf{1}^T \qquad \textbf{(3.60)}$$

(b) Expectation

$$E(\hat{N}) = \Upsilon Q[I - Q]^{-1}\mathbf{1}^T \qquad \textbf{(3.61)}$$

(c) Variance

$$Var(\hat{N}) = \Upsilon Q^2[I - Q]^{-2} + Q[I - Q]^{-1}\mathbf{1}^T \qquad \textbf{(3.62)}$$

Proof of theorem 3.2.5

(a) The pmf of phase type Geometric distribution is:

$$y(\hat{n}) = Q^n[I - Q] \qquad y(\hat{n}) = Q^nP$$

The pgf of n is given by:

$$Y(s) = \sum_{n=0}^{\infty} y(\hat{n})s^n \qquad = \sum_{n=0}^{\infty} Q^nPs^n \qquad = P[I - Qs]^{-1} \qquad \textbf{(3.63)}$$

Equation (3.63) can be transformed to a proper pgf by multiplying by Υ on the LHS and $\mathbf{1}^T$ on the RHS to become:

$$Y(\hat{s}) = \Upsilon[I - Q][I - Qs]^{-1}\mathbf{1}^T \qquad \textbf{(3.64)}$$

which is the pgf of Phase type Geometric distribution. Derivative of pgf can be used to obtain other properties of phase type Geometric distribution. The derivatives of equation (3.63) are obtained as:

$$Y'(s) = QP[I - Qs]^{-2} \qquad \textbf{(3.65)}$$

$$Y''(s) = 2Q^2P[I - Qs]^{-3} \qquad \textbf{(3.66)}$$

In general the k^{th} factorial moments is given by:

$$Y^k(s) = k!Q^kP[I - Qs]^{-(k+1)}$$

(b) $E(N) = Y'(1)$ hence equation (3.65) can be expressed as:

$$Y'(1) = QPP^{-2} \qquad Y'(1) = QP^{-1}$$

Hence:

$$E(N) = QP^{-1} \qquad \textbf{(3.67)}$$

Equation (3.67) can be wriNen as a proper expectation as:

$$E(\hat{N}) = \Upsilon QP^{-1}\mathbf{1}^T \qquad \textbf{(3.68)}$$

(c) Variance of N can be expressed as shown in (3.24) and $Y''(1)$ can be obtained from equation (3.66) as:

$$Y''(1) = 2Q^2PP^{-3} \qquad Y''(1) = 2Q^2P^{-2} \tag{3.69}$$

Hence equation (3.24) can be expressed as:

$$Var(N) = 2Q^2P^{-2} + QP^{-1} - Q^2P^{-2} = Q^2P^{-2} + QP^{-1} \tag{3.70}$$

Equation (3.70) can be rewriNen as :

$$Var(\hat{N}) = \Upsilon Q^2P^{-2} + QP^{-1}\mathbf{1}^T \tag{3.71}$$

3.3 Compound distributions of Panjer class $(a, b, 0)$ distributions with severity distributions

This section develops and applies compound phase type distributions (CPHD) in modeling secondary cancer cases for distributions of Panjer class $(a, b, 0)$. Compound distributions considered in previous researches do not in cooperate diseases with transition phases, which can be considered by compound phase type distribution(CPHD). Convoluting pgf of count distribution and pgf of severity distribution, the pgf of compound distributions can be derived.

Definition 3.3.1. *Let N be a random variable whose pgf is $Y(s)$ and $G_1, \ldots, G_N$ is a set of independent and identically distributed random variable which has a common pgf $X(s)$ which is independent from N, therefore the pgf of compound distribution is illustrated as:*

$$Z(s) = Y[X(s)] \tag{3.72}$$

3.3.1 General expression of phase type compound distributions

Theorem 3.3.2 (Compound PH-Poisson). *If the pgf of $N \sim PH - P(\Lambda)$ the compound pgf of N is:*

$$Z(s) = \Upsilon e^{\Lambda\, L_x[X(s)]-I}\,\mathbf{1}^T \tag{3.73}$$

where $L_x[X(S)]$ represents Laplace transform of the claim amount distribution for continuous severity distributions.

Proof of theorem 3.3.2

Let the pgf of the compound distribution be illustrated as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (3.18) it becomes:

$$Z(s) = Y\left[L_x[X(s)]\right] = \Upsilon e^{\Lambda\, L_x[X(s)]-I}\,\mathbf{1}^T \tag{3.74}$$

Theorem 3.3.3 (Compound PH-Negative Binomial). *If the pgf of $N \sim PH - NBin(Q, \alpha)$ the compound pgf of N becomes:*

$$Z(s) = \Upsilon\left(P\left[I - L_x[X(s)][I-P]\right]^{-1}\right)^{\alpha} 1^T \tag{3.75}$$

where $L_x[X(s)]$ is as the Laplace transform.

Proof of theorem 3.3.3

Let the pgf of the compound distribution be illustrated as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (3.35) it becomes:

$$Z(s) = Y\, L_x[X(s)] \qquad = \Upsilon\left(P\left(I - L_x[X(s)][I-P]\right)^{-1}\right)^{\alpha} 1^T \tag{3.76}$$

Theorem 3.3.4 (Compound PH-Binomial). *If the pgf of $N \sim PH - Bin(Q, \alpha)$ the the compound pgf of N becomes:*

$$Z(s) = \Upsilon\left[P\, L_x[X(s)] + [I-P]\right]^{\alpha} 1^T \tag{3.77}$$

where $L_x[X(s)]$ is as the Laplace transform.

Proof theorem 3.3.4

Let the pgf of the compound distribution be illustrated as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (3.48) it becomes:

$$Z(s) = Y[L_x[X(s)]] \qquad = \Upsilon\left(P\, L_x[X(s)] + [I-P]\right)^{\alpha} 1^T \tag{3.78}$$

Theorem 3.3.5 (Compound PH Geometric). *If the pgf of $N \sim PH - Geo(Q)$ the the compound pgf of N becomes:*

$$Z(s) = \Upsilon\left[I - L_x[X(s)][I-P]\right]^{-1} P 1^T \tag{3.79}$$

where $L_x[X(s)]$ is as the Laplace transform.

Proof of theorem 3.3.5

Let the pgf of the compound distribution be expressed as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (3.63) it becomes:

$$Z(s) = Y[L_x[X(s)]] \qquad = \Upsilon\left(I - L_x[X(s)][I-P]\right)^{-1} P 1^T \tag{3.80}$$

3.3.2 Laplace transform and probability generating function of severity distributions

The continuous severity distributions taken into consideration in this study are; Weibull, Pareto and Generalized Pareto. Most continuous distributions do not have probability generating function hence their Laplace transforms are derived and replaced in equation (3.74), (3.76), (3.78) and (3.80) to get the pgf of the compound distributions.

(i) Weibull distribution

The Laplace transform of Weibull distribution can be derived as:

$$L_x X(s) = E[e^{-sx}]$$

$$L_x X(s) = \int_0^\infty e^{-sx} \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(\frac{x}{\alpha})^\beta} dx$$

$$L_x X(s) = \frac{\beta}{\alpha} \frac{\Gamma(\beta)}{[s\alpha + (\frac{x}{\alpha})^{\beta-1}]^\beta} \quad \textbf{(3.81)}$$

(ii) Pareto distribution

The Laplace transform of Pareto distribution can be derived as:

$$L_x X(s) = E[e^{-sx}]$$

$$L_x X(s) = \alpha\beta^\alpha \int_0^\infty \frac{e^{-sx}}{(x+\beta)^{\alpha+1}} dx$$

$$L_x X(s) = \frac{\alpha}{\beta} \sum_{k=0}^{\infty} (-1)^k \frac{\Gamma(\alpha+k)}{k!\,\Gamma(\alpha)} \frac{k!}{\beta^{2k+1}}$$

$$L_x X(s) = \sum_{k=0}^{\infty} (-1)^k \frac{\alpha\,\Gamma(\alpha+k)}{\Gamma(\alpha)\,\beta^{2k+2}} \quad \textbf{(3.82)}$$

(iii) Generalized Pareto distribution

The Laplace transform of Generalized Pareto distribution can be derived as:

$$L_x X(s) = E[e^{-sx}]$$

$$L_x X(s) = \int_0^\infty e^{-sx} \frac{x^{\alpha-1}}{\beta(\alpha,\gamma)(x+\lambda)^{\alpha+\gamma}} dx$$

$$L_x X(s) = \frac{1}{\lambda^\gamma \beta(\alpha,\gamma)} \int_0^\infty x^\alpha e^{-sx} \sum_{k=0}^{\infty} \binom{\alpha+\gamma}{k} \left(\frac{x}{\lambda}\right)^k dx$$

$$= \frac{1}{\lambda^\gamma \beta(\alpha,\gamma)} \sum_{k=0}^{\infty} \frac{\binom{-(\alpha+\gamma)}{k}}{\lambda^k} \int_0^\infty x^{\gamma+k+1-1} e^{-sx} dx$$

$$L_x X(s) = \frac{1}{\lambda^\gamma \beta(\alpha,\gamma)} \sum_{k=0}^{\infty} \frac{\binom{-(\alpha+\gamma)}{k}}{\lambda^k} \frac{\Gamma(\alpha+k)}{s^{\alpha+k}} \quad \textbf{(3.83)}$$

The discrete distributions considered in this study are one parameter Poisson Lindley and two parameter Poisson Lindley distributions which have probability generating functions expressed as:

(i) One parameter Poisson Lindley distribution

The probability generation function of one parameter Poisson Lindley distribution can be derived as:

$$X(s) = \int_0^\infty e^{\lambda(1-s)} \frac{\theta^2}{I+\theta}(1+\lambda)e^{-\theta\lambda}\, d\lambda$$

$$X(s) = \frac{\theta^2}{I+\theta}\left[\int_0^\infty \lambda e^{-\lambda(1+\theta-s)} d\lambda + \int_0^\infty e^{-\lambda(\theta+1-s)} d\lambda\right]$$

$$X(s) = \frac{\theta^2}{1+\theta}\left[\frac{\theta+2-s}{[\theta+1-s]^2}\right] \qquad \textbf{(3.84)}$$

(ii) Two parameter Poisson Lindley distribution

The probability generation function of two parameter Poisson Lindley distribution can be derived as:

$$X(s) = \frac{\theta^2}{(\theta+1)^2}\sum_{x=0}^{\infty}\left[\frac{s}{(\theta+1)}\right]^x + \frac{\theta^2}{(\theta+1)^2(\alpha\theta+1)}\sum_{x=0}^{\infty}(\alpha+x)\left[\frac{s}{\theta+1}\right]^x$$

$$X(s) = \frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2} \qquad \textbf{(3.85)}$$

3.3.3 Compound phase type distributions probability generating functions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (3.74) the pgf of the compound distributions of PH-Poisson with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH Poisson-Weibull	$Z(s) = \boldsymbol{\Upsilon} e^{\Lambda\left(\frac{\beta}{\alpha}\frac{\Gamma(\beta)}{[s\alpha+(\frac{x}{\alpha})^\beta-1]^\beta}-1\right)} \mathbf{1}^T$
PH Poisson-Pareto	$Z(s) = \boldsymbol{\Upsilon} e^{\Lambda\left(\sum_{k=0}^{\infty}(-1)^k \frac{\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}-1\right)} \mathbf{1}^T$
PH Poisson-Gen Pareto	$Z(s) = \boldsymbol{\Upsilon} e^{\Lambda\left(\frac{1}{\lambda^\gamma \beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\left[\frac{-(\alpha+\gamma)}{\lambda^k}\frac{\Gamma(\alpha+k)}{s^{\alpha+k}}\right]-1\right)} \mathbf{1}^T$
PH Poisson-OPPL	$Z(s) = \boldsymbol{\Upsilon} e^{\Lambda\left(\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}-1\right)} \mathbf{1}^T$
PH Poisson-TPPL	$Z(s) = \boldsymbol{\Upsilon} e^{\Lambda\left(\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}-1\right)} \mathbf{1}^T$

Table 3.1. Compound phase type Poisson distributions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (3.76) the pgf of the compound distributions of PH-Negative Binomial with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH Neg Binomial-Weibull	$Z(s) = \text{Ɏ}, P[I - \frac{\beta}{\alpha}\frac{\Gamma(\beta)}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}} (I-P)]^{-1}, \alpha\, \mathbf{1}^T$
PH Neg Binomial-Pareto	$Z(s) = \text{Ɏ}, P[I - \sum_{k=0}^{\infty}(-1)^k \frac{\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}} (I-P)]^{-1}, \alpha\, \mathbf{1}^T$
PH Neg Binomial-Gen Pareto	$Z(s) = \text{Ɏ}, P[I - \frac{1}{\lambda^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\lambda^k}\frac{\Gamma(\alpha+k)}{s^{\alpha+k}} (I-P)]^{-1}, \alpha\, \mathbf{1}^T$
PH Neg Binomial-OPPL	$Z(s) = \text{Ɏ}, P[I - \frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2} (I-P)]^{-1}, \alpha\, \mathbf{1}^T$
PH Neg Binomial-TPPL	$Z(s) = \text{Ɏ}, P[I - \frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2} (I-P)]^{-1}\, \mathbf{1}^T$

Table 3.2. Compound phase type Negative Binomial distributions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (3.78) the pgf of the compound distributions of PH-Binomial with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH Binomial-Weibull	$Z(s) = \text{Ɏ}, P \frac{\beta}{\alpha}\frac{\Gamma(\beta)}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}} + (I-P), \alpha\, \mathbf{1}^T$
PH Binomial-Pareto	$Z(s) = \text{Ɏ}, P \sum_{k=0}^{\infty}(-1)^k \frac{\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}} + (I-P), \alpha\, \mathbf{1}^T$
PH Binomial-Gen Pareto	$Z(s) = \text{Ɏ}, P \frac{1}{\lambda^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\lambda^k}\frac{\Gamma(\alpha+k)}{s^{\alpha+k}} + (I-P), \alpha\, \mathbf{1}^T$
PH Binomial-OPPL	$Z(s) = \text{Ɏ}, P \frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2} + (I-P), \alpha\, \mathbf{1}^T$
PH Binomial-TPPL	$Z(s) = \text{Ɏ}, P \frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2} + (I-P)\, \mathbf{1}^T$

Table 3.3. Compound phase type Binomial distributions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (3.80) the pgf of the compound distributions of PH-Geometric with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH Geometric-Weibull	$Z(s) = \text{Ɏ}, I - \frac{\beta}{\alpha}\frac{\Gamma(\beta)}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}} (I-P), ^{-1} P\mathbf{1}^T$
PH Geometric-Pareto	$Z(s) = \text{Ɏ}, I - \sum_{k=0}^{\infty}(-1)^k \frac{\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}} (I-P), ^{-1} P\mathbf{1}^T$
PH Geometric-Gen Pareto	$Z(s) = \text{Ɏ} \; I - \frac{1}{\lambda^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\lambda^k}\frac{\Gamma(\alpha+k)}{s^{\alpha+k}} (I-P) \; P\mathbf{1}^T$
PH Geometric-OPPL	$Z(s) = \text{Ɏ} \; I - \frac{}{1+\theta}\frac{}{[\theta+1-s]^2} (I-P) \; P\mathbf{1}$
PH Geometric-TPPL	$Z(s) = \text{Ɏ}, I - \frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2} (I-P), ^{-1} P\mathbf{1}^T$

Table 3.4. Compound phase type Geometric distributions

3.4 Phase type Panjer recursion formula for class $(a, b, 0)$

Phase type Panjer recursion formula is obtained from the pgf of the phase type distributions of aggregate losses, claim count distribution and claim amount distributions. Let $Z(s)$, $Y(s)$ and $X(s)$

be $M * M$ matrix representing the phase type aggregate loss, phase type claim count distribution pgf and phase type claim amount distribution pgf respectively.

Theorem 3.4.1. *Phase type compound distribution of class $(a, b, 0)$ recursively If the distribution of N belongs to Panjer class $(a, b, 0)$ therefore it should satisfies the following equation:*

$$Z(j) = \sum_{x=1}^{\infty} y(x) Z(j - x) \left(A + B\frac{x}{j}\right)\left[I - Ay(0)\right]^{-1}$$

where $Z(j)$ is a $M * M$ matrix and $Z(0)$ is the initial matrix.

Proof of theorem 3.4.1

The pgf of phase type aggregate loss distribution is expressed as:

$$Z(s) = E[s^{S_N}] \qquad = \sum_{j=0}^{\infty} Z(j)s^j$$

The pgf of phase type frequency distribution is expressed as:

$$Y(s) = E[s^N] \qquad = \sum_{n=0}^{\infty} P_n s^n$$

The pgf of severity distribution is expressed as:

$$X(s) = E[s^{X_i}] \qquad = \sum_{x=0}^{\infty} y(x)s^x$$

Phase type aggregate loss distribution can be represented as a convolution of the pgf of phase type claim count distribution and pgf of claim amount distribution as:

$$Z(s) = Y[X(s)] \tag{3.86}$$

Equation's (3.86) first derivative is:

$$Z'(s) = \{Y'[X'(s)]\}X'(s) \tag{3.87}$$

The pgf of phase type claim count distribution is:

$$Y(s) = \sum_{n=0}^{\infty} P_n s^n \tag{3.88}$$

Combining equation (3.86)and equation (3.88) it becomes:

$$Z(s) = \sum_{n=0}^{\infty} P_n [X(s)]^n \tag{3.89}$$

The first derivative of the pgf of the phase type aggregate loss distribution as expressed in equation (3.89) is:

$$Z'(s) = \sum_{n=0}^{\infty} nP_n[X(s)]^{n-1}X'(s) \tag{3.90}$$

Phase type Panjer's model is expressed as :

$$P_n = \quad A + \frac{B}{n} \quad P_{n-1} \qquad\qquad n = 1, 2, 3, \ldots \tag{3.91}$$

Multiplying all through by n equation (3.91) becomes:

$$nP_n = \quad nA + B \quad P_{n-1} \qquad\qquad n = 1, 2, 3, \ldots \tag{3.92}$$

Multiplying equation (3.92) by $[X(s)]^{n-1}X'(s)$ and adding the result over n it results to :

$$\sum_{n=1}^{\infty} nP_n[X(s)]^{n-1}X'(s) = \sum_{n=1}^{\infty} (nA + B)P_{n-1}[X(s)]^{n-1}X'(s) \tag{3.93}$$

Merging equation (3.90) and equation (3.93) it becomes :

$$Z'(s) = \sum_{n=1}^{\infty} (n-1)AP_{n-1}[X(s)]^{n-1}X'(s) + (A+B)\sum_{n=1}^{\infty} P_{n-1}[X(s)]^{n-1}X'(s) \tag{3.94}$$

By definition the pgf of phase type aggregate loss distribution is:

$$Z(s) = \sum_{j=0}^{\infty} Z(j)s^j \tag{3.95}$$

The first derivative of equation (3.95) is :

$$Z'(s) = \sum_{j=1}^{\infty} jZ(j)s^{j-1} \tag{3.96}$$

By definition the pgf of severity distribution is:

$$X(s) = \sum_{x=0}^{\infty} y(x)s^x \tag{3.97}$$

Equation's (3.97) first derivative can be expressed as:

$$X'(s) = \sum_{x=1}^{\infty} xy(x)s^{x-1} \tag{3.98}$$

Merging equation (3.94) and equation (3.96) it becomes:

$$\sum_{j=1}^{\infty} jZ(j)s^{j-1} = A \sum_{n=1}^{\infty}(n-1)P_{n-1}[X(s)]^{n-1}X'(s)\} + (A+B) \sum_{n=1}^{\infty}P_{n-1}[X(s)]^{n-1}X'(s)\} \qquad \textbf{(3.99)}$$

Apart from equation (3.95) $Z(s)$ can be shown as:

$$Z(s) = \sum_{n=1}^{\infty} P_{n-1}[X(s)]^{n-1}X(s)' \qquad \textbf{(3.100)}$$

Hence the first derivative of equation (3.100) ;

$$Z'(s) = \sum_{n=2}^{\infty}(n-1)P_{n-1}[X(s)]^{n-2}X'(s) \qquad \textbf{(3.101)}$$

Merging equation (3.100), (3.101) and equation (3.99) it becomes:

$$\sum_{j=1}^{\infty} jZ(j)s^{j-1} = A\, X(s)Z'(s) + (A+B)Z(s)X'(s)$$
$$= A\sum_{x=0}^{\infty} y(x)s^{x} \sum_{j=1}^{\infty} jZ(j)s^{j-1} + (A+B)\sum_{j=0}^{\infty} Z(j)s^{j} \sum_{x=1}^{\infty} xy(x)s^{x-1} \qquad \textbf{(3.102)}$$

Multiplying equation (3.102) by equation (3.90) it becomes :

$$\sum_{j=1}^{\infty} jZ(j)s^{j} = A \sum_{x=0}^{\infty} y(x)s^{x} \sum_{j=1}^{\infty} jZ(j)s^{j} + (A+B)\sum_{j=0}^{\infty} Z(j)s^{j} \sum_{x=1}^{\infty} xy(x)s^{x} \qquad \textbf{(3.103)}$$

But we know:

$$\sum_{j=1}^{\infty} jZ(j)s^{j} = \sum_{j=x+1}^{\infty} (j-x)Z(j-x)s^{j-x} \qquad \textbf{(3.104)}$$

and

$$\sum_{j=0}^{\infty} Z(j)s^{j} = \sum_{j=x}^{\infty} Z(j-x)s^{j-x} \qquad \textbf{(3.105)}$$

Combining equation (3.103), (3.104) and (3.105) it becomes :

$$\sum_{j=1}^{\infty} jZ(j)s^{j} = A\sum_{x=0}^{\infty}\sum_{j=x+1}^{\infty} (j-x)y(x)Z(j-x)s^{j} + (A+B)\sum_{j=x}^{\infty} Z(j-x)s^{j-x} \sum_{x=1}^{\infty} xy(x)s^{x}$$
$$= A\sum_{x=0}^{\infty}\sum_{j=x+1}^{\infty} (j-x)y(x)Z(j-x)s_{j} + (A+B)\sum_{j=1}^{\infty}\sum_{x=1}^{\infty} xy(x)Z(j-x)s_{j}$$
$$= \sum_{j=1}^{\infty}\{A \sum_{x=0}^{\infty} (j-x)y(x)Z(j-x)\}s_{j} + \sum_{j=1}^{\infty}\{(A+B) \sum_{x=1}^{\infty} xy(x)Z(j-x)\}s_{j} \qquad \textbf{(3.106)}$$

Combining the coefficients of s^j in equation (3.106) it becomes :

$$\begin{aligned} jZ(j) &= A\sum_{x=0}^{\infty}(j-x)y(x)Z(j-x) + (A+B)\sum_{x=1}^{\infty}xy(x)Z(j-x) \\ &= Ajy(0)Z(j) + A\sum_{x=1}^{\infty}(j-x)y(x)Z(j-x) + (A+B)\sum_{x=1}^{\infty}xy(x)Z(j-x) \end{aligned} \quad (3.107)$$

Factoring out $I - Ay(0)$ in equation (3.107) it becomes:

$$\begin{aligned} \left[I - Ay(0)\right] jZ(j) &= \sum_{x=1}^{\infty}\left[A(j-x) + (A+B)x\right] y(x)Z(j-x) \\ \left[I - Ay(0)\right] jZ(j) &= \sum_{x=1}^{\infty}(Aj + Bx)\, y(x)Z(j-x) \end{aligned} \quad (3.108)$$

The recursive form of the compound distribution is found by dividing equation (3.108) by $\left[I - Ay(0)\right] j$ to get:

$$Z(j) = \left[\sum_{x=1}^{\infty}(Aj+Bx)\, y(x)Z(j-x)\right]\left[\left(I - Ay(0)\right) j\right]^{-1} \quad (3.109)$$

Equation (3.109) can be rearranged to:

$$Z(j) = \sum_{x=1}^{\infty} y(x)Z(j-x)\left(A + B\frac{x}{j}\right)\left[I - Ay(0)\right]^{-1} \quad (3.110)$$

$\mathbf{Z}(j)$ can be expressed as a row vector as $\boldsymbol{\Upsilon}Z(j)$ hence equation (3.110) becomes:

$$\mathbf{Z}(j) = \sum_{x=1}^{\infty} y(x)\mathbf{Z}(j-x)\left(A + B\frac{x}{j}\right)\left[I - Ay(0)\right]^{-1} \quad (3.111)$$

The initial aggregate loss probability matrix is expressed as:

$$Z(0) = \sum_{n=0}^{\infty}\left[y(0)\right]^n P_n \quad (3.112)$$

$Z(j)$ can be expressed as a probability as:

$$z(j) = \boldsymbol{\Upsilon}Z(j)\mathbf{1}^T$$

3.5 Phase type distributions of class (a,b,1) and its aggregate loss distribution recursively

3.5.1 Phase type distributions of class $(a, b, 1)$ using iteration technique

Theorem 3.5.1 (Phase type Panjer class $(a, b, 1)$ distributions). *The distributions arising from phase type Panjer recursive model*

$$P_n = \left(A + \frac{B}{n}\right) P_{n-1} \quad for\ n = 2, 3, 4, \ldots \qquad \textbf{(3.113)}$$

are:

(i) Phase type Zero-truncated Poisson distribution when $A = 0, B > 0$.

(ii) Phase type Shifted (zero-truncated) geometric distribution when $A > 0, B = 0$.

(iii) Phase type Zero-truncated Negative Binomial distribution when $A > 0, A + B > 0$.

Proof of theorem 3.5.1

(i) **When** $A = 0, B > 0$

Equation (3.113) becomes:

$$P_n = \frac{B}{n} P_{n-1} \qquad \textbf{(3.114)}$$

$n = 2$;

$$P_2 = \frac{B}{2} P_1$$

$n = 3$;

$$P_3 = \frac{B}{3} P_2 = \frac{B}{3} * \frac{B}{2} P_1 = \frac{B^2}{3!} P_1$$

$n = k$;

$$P_k = \frac{B^{k-1}}{k!} P_1 \qquad k = 2, 3, 4 \ldots$$

P_k can be expressed as:

$$\sum_{k=1}^{\infty} P_k = I \qquad P_1 + \sum_{k=2}^{\infty} P_k = I \qquad P_1 + \sum_{k=2}^{\infty} \frac{B^{k-1}}{k!} P_1 = I$$

$$P_1 \left(I + \sum_{k=2}^{\infty} \frac{B^{k-1}}{k!} \right) = I \qquad P_1 B^{-1} \sum_{k=1}^{\infty} \frac{B^k}{k!} = I \qquad P_1 B^{-1} (e^B - I) = I$$

$$P_1 = B(e^B - I)^{-1}$$

Therefore:

$$P_k = \frac{B^{k-1}}{k!} * B(e^B - I)^{-1} \qquad P_k = [B^k (Bk!)^{-1}] * [(B)(e^B - I)^{-1}] \tag{3.115}$$

$$P_k = [B^k][k!(e^B - I)]^{-1} \qquad k = 0, 1, 2, \ldots \tag{3.116}$$

This is phase type Zero-truncated Poisson distribution with parameter $B > 0$.

Theorem 3.5.2 (Properties of phase type Zero-truncated Poisson distribution). *Letting A from equation (3.116) be* 0 *and* $B = \Lambda$ *the properties of phase type Zero-truncated Poisson distribution are expressed as:*

(a) pgf

$$X(\hat{s}) = \Upsilon[\Lambda e^{\Lambda s}][e^{\Lambda} - I] - 1\mathbb{1}^T \tag{3.117}$$

(b) Expectation

$$E(\hat{N}) = \Upsilon \Lambda [I - e^{-\Lambda}]^{-1} \mathbb{1}^T \tag{3.118}$$

(c) Variance

$$Var(\hat{N}) = \Lambda [I - e^{-\Lambda}]^{-1} \tag{3.119}$$

Proof of theorem 3.5.2

(a) The pmf of phase type Zero truncated Poisson distribution can be shown as:

$$f(n) = \Lambda^n [(e^{\Lambda} - I) n!]^{-1} \quad n = 1, 2, 3, \ldots$$

The pgf of n is given by:

$$X(s) = \sum_{n=1}^{\infty} f(n) s^n \qquad = \sum_{n=1}^{\infty} \Lambda^n [(e^{\Lambda} - I) n!]^{-1} s^n \qquad = [I(e^{\Lambda} - I)^{-1}] * e^{\Lambda s} - I$$

$$X(s) = [e^{\Lambda s} - I][e^{\Lambda} - I]^{-1} \tag{3.120}$$

Equation (3.120) can be transformed to a proper pgf by multiplying by Υ on the LHS and $\mathbb{1}^T$ on the RHS to become:

$$X(\hat{s}) = \Upsilon [e^{\Lambda s} - I][e^{\Lambda} - I]^{-1} \mathbb{1}^T \tag{3.121}$$

Derivative of pgf can be used to obtain other properties of phase type Poisson distribution. The derivatives of equation (3.120) are obtained as:

$$X'(s) = [\Lambda e^{\Lambda s}][e^{\Lambda} - I]^{-1} \tag{3.122}$$

$$X''(s) = [\Lambda^2 e^{\Lambda s}][e^{\Lambda} - I]^{-1} \tag{3.123}$$

In general the k^{th} factorial moments is given by:

$$X^k(s) = [\Lambda^k e^{\Lambda s}][e^{\Lambda} - I]^{-1}$$

(b) $E(N) = X'(1)$ hence equation (3.122) can be expressed as:

$$X'(s) = [\Lambda e^{\Lambda s}][e^{\Lambda} - I]^{-1} \qquad X'(1) = [\Lambda e^{\Lambda}][e^{\Lambda} - I]^{-1} \qquad X(1) = \Lambda[I - e^{-\Lambda}]^{-1}$$

Hence:

$$E(N) = \Lambda[I - e^{-\Lambda}]^{-1} \qquad E(\hat{N}) = \text{Ɏ}\Lambda[I - e^{-\Lambda}]^{-1}1^{T} \tag{3.124}$$

(c) Variance of N can be expressed as:

$$Var(N) = X''(1) + X'(1) - [X'(1)]^2 \tag{3.125}$$

$X''(1)$ can be obtained from equation (3.123) as:

$$X''(s) = [\Lambda^2 e^{\Lambda}][e^{\Lambda} - I]^{-1} \qquad X''(s) = \Lambda^2[I - e^{-\Lambda}]^{-1} \tag{3.126}$$

Hence equation (3.126) can be illustrated as:

$$Var(N) = \Lambda^2[I - e^{-\Lambda}]^{-1} + \Lambda[I - e^{-\Lambda}]^{-1} - \Lambda^2[I - e^{-\Lambda}]^{-2} \qquad = \Lambda[I - e^{-\Lambda}]^{-1}$$

$$Var(N) = \text{Ɏ}\Lambda[I - e^{-\Lambda}]^{-1}1^{T} \tag{3.127}$$

(ii) **When** $A > 0, B = 0$

Equation (3.113) becomes:

$$P_n = AP_{n-1} \tag{3.128}$$

$n = 2$**;**

$$P_2 = AP_1$$

$n = 3$**;**

$$P_3 = AP_2 \qquad = A * AP_1 \qquad = A^2 P_1$$

$n = k$**;**

$$P_k = A^{k-1}P_1 \qquad k = 2, 3, 4\ldots$$

P_k can be expressed as:

$$\sum_{k=1}^{\infty} P_k = I \qquad P_1 + \sum_{k=2}^{\infty} P_k = 1 \qquad P_1 + \sum_{k=2}^{\infty} A^{k-1}P_1 = I$$

$$P_1 \;\; I + \sum_{k=2}^{\infty} A^{k-1} \;\; = I \qquad P_1 \sum_{k=1}^{\infty} A^{k-1} = I \qquad P_1(I - A)^{-1} = I$$

$$P_1 = (I - A)$$

Therefore:

$$P_k = A^{k-1}(I - A) \tag{3.129}$$

This is a phase type Zero-truncated geometric distribution with parameter $0 < I - A < I$.

Theorem 3.5.3 (**Properties of phase type Zero-truncated geometric distribution**). *Letting A from equation (3.129) be > Q and B = 0 the properties of phase type Zero-truncated geometric distribution are expressed as:*

(a) pgf

$$X(\hat{s}) = \Upsilon P[Q - Q^2 s]^{-1}\mathbb{1}^T \tag{3.130}$$

(b) Expectation

$$E(\hat{N}) = \Upsilon P^{-1}\mathbb{1}^T \tag{3.131}$$

(c) Variance

$$Var(\hat{N}) = \Upsilon I - P^{-1}\mathbb{1}^T \tag{3.132}$$

Proof of theorem 3.5.3

(a) The pmf of phase type Geometric distribution is:

$$f(n) = P[I - P]^{n-1} \qquad f(n) = PQ^{n-1}$$

The pgf of n is given by:

$$X(s) = \sum_{n=0}^{\infty} f(n)s^n = \sum_{n=0}^{\infty} P[I - P]^{n-1}s^n$$

$$X(s) = \frac{P}{I - P}\sum_{n=0}^{\infty}[(I - P)s]^n = \frac{P}{I - P} * \frac{I}{I - (I - P)s} \tag{3.133}$$

$$X(s) = \Upsilon P[Q - Q^2 s]^{-1}\mathbb{1}^T \tag{3.134}$$

which is the pgf of Phase type Zero truncated Geometric distribution. The derivatives of equation (3.134) are obtained as:

$$X'(s) = Q^2 P[Q - Q^2 s]^{-2} \tag{3.135}$$

$$X''(s) = Q^4 P[Q - Q^2 s]^{-3} \tag{3.136}$$

In general the k^{th} factorial moments is given by:

$$X^k(s) = Q^{2k}P[Q - Q^2 s]^{-k-1}$$

(b) $E(N) = X'(1)$ hence equation (3.135) can be expressed as:

$$X'(1) = Q^2P[Q - Q^2]^{-2} \qquad = P[I - Q]^{-2} \qquad = P^{-1}$$

Hence:

$$E(N) = P^{-1} \qquad E(\hat{N}) = \Upsilon P^{-1} 1^T \tag{3.137}$$

(c) Variance of N can be expressed as shown in (3.125) and $G''(1)$ can be obtained from equation (3.136) as:

$$X''(1) = Q^4P[Q - Q^2]^{-3} \qquad = QP[I - Q]^{-3} \qquad = QP^{-2} \tag{3.138}$$

Hence equation (3.125) can be expressed as:

$$Var(N) = Q[I - Q]^{-2} + [I - Q]^{-1} - [I - Q]^{-2} \qquad = Q[I - Q]^{-2} + I[I - Q]^{-1} - I[I - Q]^{-2}$$

$$Var(N) = I - P^{-1} \qquad Var(\hat{N}) = \Upsilon I - P^{-1} 1^T \tag{3.139}$$

(iii) **Where** $A > 0, A + B > 0$

Equation (3.113) becomes:

$$P_n = \frac{nA + B}{n} P_{n-1} \qquad n = 2, 3, 4, \ldots \tag{3.140}$$

$n = 2$;

$$P_2 = \frac{2A + B}{2} P_1$$

$n = 3$;

$$P_3 = \frac{3A + B}{3} P_2 \qquad = \frac{3A + B}{3} * \frac{2A + B}{2} P_1 \qquad = (3A + B)(2A + B)\frac{P_1}{3!} \tag{3.141}$$

$n = k$;

$$P_k = (kA + B)[(k-1)A + B]\ldots(3A + B)(2A + B)\frac{P_1}{k!}$$

$$= A^{k-1}\left[kI + BA^{-1}\right]\left[(k-1)I + BA^{-1}\right]\cdots\left[3I + BA^{-1}\right]\left[2I + BA^{-1}\right]\frac{P_1}{k!}$$

$$= A^k(A + B)^{-1}\left[BA^{-1} + kI\right]\left[BA^{-1} + (k-1)I\right]\ldots\left[BA^{-1} + 3I\right]\left[BA^{-1} + 2I\right]\left[BA^{-1} + I\right]\frac{P_1}{k!}$$

$$= A^k(A + B)^{-1}\binom{BA^{-1} + kI}{kI} P_1 \tag{3.142}$$

P_k can be expressed as:

$$\sum_{k=1}^{\infty} P_k = I$$

$$P_1 + \sum_{k=2}^{\infty} A^k (A+B)^{-1} \binom{BA^{-1}+kI}{kI} P_1 = I$$

$$P_1 \sum_{k=1}^{\infty} (A^k)(A+B)^{-1} \binom{BA^{-1}+kI}{kI} = I$$

$$P_1 (A+B)^{-1} \sum_{k=1}^{\infty} A^k \binom{BA^{-1}+I+kI-I}{kI} = I$$

$$P_1 (A+B)^{-1} \sum_{k=1}^{\infty} A^k (-1)^k \binom{-[(BA)A^{-1}+I]}{kI} = I$$

$$P_1 (A+B)^{-1} \sum_{k=1}^{\infty} \binom{-[(A+B)A^{-1}+I]}{kI} (-A)^k = I$$

$$P_1 (A+B)^{-1} \left[(I-A)^{-[(A+B)A^{-1}]} - I\right] = I$$

$$P_1 (A+B)^{-1} = [(I-A)^{-[(A+B)A^{-1}]} - I]^{-1}$$

$$P_1 = (A+B)[(I-A)^{-[(A+B)A^{-1}]} - I]^{-1} \quad k = 1, 2, 3, \ldots \qquad \textbf{(3.143)}$$

Therefore:

$$P_k = A^k (A+B)^{-1} \binom{BA^{-1}+kI}{kI} (A+B)[(I-A)^{-[(A+B)A^{-1}]} - I]^{-1}$$

$$P_k = \binom{-(A+B)A^{-1}}{kI} (-A)^k \left[(I-A)^{-[(A+B)A^{-1}]} - I\right]^{-1} \qquad \textbf{(3.144)}$$

This is Zero-truncated Negative Binomial distribution with parameter $(A+B)(A)^{-1} > 0$ $0 < I - A < I$

Theorem 3.5.4 (Properties of phase type zero-truncated Negative Binomial distribution). *Letting A from equation (3.144) be* $\beta\,[I + \beta I]^{-1}$ *and* $B = [(r-1)\beta\,][I + \beta I]^{-1}$ *the properties of phase type Zero-truncated Negative Binomial distribution are expressed as:*

(a) pgf

$$X(s) = \Upsilon \left\{[I - \beta I(s-1)]^{-r} - [I+\beta I]^{-r}\right\} \left\{I - (I+\beta I)^{-r}\right\}^{-1} \mathbf{1}^T \qquad \textit{(3.145)}$$

(b) Expectation

$$E(N) = \Upsilon\, r\beta \left\{I - (I+\beta I)^{-r}\right\}^{-1} \mathbf{1}^T \qquad \textit{(3.146)}$$

(c) Variance

$$Var(N) = \Upsilon \left\{ r\beta[(I+\beta I) - (1+\beta+r\beta)(I+\beta)^{-r}] \right\} \left\{ [I-(I+\beta I)^{-r}]^{-2} \right\}^{-1} \mathbf{1}^T \tag{3.147}$$

Proof of theorem 3.5.4

(a) Let the $X(s) = \sum_{n=1}^{\infty} f(n)s^n$ hence:

$$= \sum_{n=1}^{\infty} s^n \; r(r+1)\ldots(r+n-1) \left\{ n![(I+\beta I)^r - I] \right\}^{-1} \left\{ \beta(I+\beta I)^{-1} \right\}^n$$

$$= \left[(I+\beta I)^r - I\right]^{-1} \sum_{n=1}^{\infty} s^n \left[\frac{r(r+1)\ldots(r+n-1)}{n!}\right] \left[\beta I(I+\beta I)^{-1}\right]^n$$

$$X(s) = \Upsilon \left\{ [I-\beta I(s-1)]^{-r} - [I+\beta I]^{-r} \right\} \left\{ I-(I+\beta I)^{-r} \right\}^{-1} \mathbf{1}^T \tag{3.148}$$

(b) The expectation of Zero-truncated Negative Binomial distribution can be derived from equation (3.148) based on the fact that $E(N) = G'(1)$ hence:

$$X(s) = r\beta s \left\{ I-(I+\beta I)^{-r} \right\}^{-1}$$

$$E(N) = \Upsilon \, r\beta \left\{ I-(I+\beta I)^{-r} \right\}^{-1} \mathbf{1}^T \tag{3.149}$$

(c) The variance of Zero-truncated Negative Binomial distribution can be derived as:

$$Var(N) = E(N^2) - [E(N)]^2$$

$$Var(N) = \Upsilon \left\{ r\beta[(I+\beta I) - (1+\beta+r\beta)(I+\beta)^{-r}] \right\} \left\{ [I-(I+\beta I)^{-r}]^{-2} \right\}^{-1} \mathbf{1}^T \tag{3.150}$$

3.6 Compound distributions of Panjer class $(a, b, 1)$ distributions with severity distributions

Compound phase type distributions (CPHD) used in modeling secondary cancer cases for distributions satisfying of Panjer class $(a, b, 1)$ are developed in this section. Distributions of Panjer class $(a, b, 1)$ are zero truncated hence they do not factor in zero claim count which is the nature of real claim count data. Convolution of probability generating functions of claim count distributions and probability generating function of claim amount distributions can be used derive probability generating function of compound distributions as shown in definition (3.3.1).

3.6.1 General expression of phase type compound distributions

Theorem 3.6.1 (Compound PH Zero-truncated Poisson distribution). *If the pgf of $N \sim PH - ZTP(\Lambda)$ the compound pgf of N is:*

$$Z(s) = \Upsilon \left[\Lambda e^{\Lambda L_x[X(s)]} \right] \left[e^{\Lambda} - I \right]^{-1} \mathbf{1}^T \tag{3.151}$$

where $L_x[X(s)]$ is as the Laplace transform.

Proof of theorem 3.6.1

Let the pgf of the compound distribution be expressed as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (3.117) it becomes:

$$\begin{aligned} Z(s) &= Y\left[L_x[X(s)]\right] \\ &= \Upsilon\left[\Lambda e^{\Lambda L_x[X(s)]}\right]\left[e^{\Lambda} - I\right]^{-1}\mathbf{1}^T \end{aligned} \tag{3.152}$$

Theorem 3.6.2 (Compound PH Zero-truncated Geometric distribution). *If the pgf of $N \sim PH-ZT\ Geo(Q)$ the the compound pgf of N is:*

$$Z(s) = \Upsilon P\left[Q - Q^2 L_x[X(s)]\right]^{-1}\mathbf{1}^T \tag{3.153}$$

where $L_x[X(s)]$ is as the Laplace transform.

Proof of theorem 3.6.2

Let the pgf of the compound distribution be expressed as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (3.130) it becomes:

$$\begin{aligned} Z(s) &= Y\left[L_x[X(s)]\right] \\ &= \Upsilon P\left[Q - Q^2 L_x[X(s)]\right]^{-1}\mathbf{1}^T \end{aligned} \tag{3.154}$$

Theorem 3.6.3 (Compound PH zero-truncated Negative Binomial distribution). *If the pgf of $N \sim PH-ZTNeg\ Bin(\beta, r)$ the compound pgf of N is:*

$$Z(s) = \Upsilon\left\{[I - \beta I(L_x[X(s)] - 1)]^{-r} - [I + \beta I]^{-r}\right\}\left\{I - (I + \beta I)^{-r}\right\}^{-1}\mathbf{1}^T \tag{3.155}$$

where $L_x[X(s)]$ is as the Laplace transform.

Proof of theorem 3.6.3

Let the pgf of the compound distribution be expressed as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (3.145) it becomes:

$$\begin{aligned} Z(s) &= Y\left[L_x[X(s)]\right] \\ &= \Upsilon\left\{[I - \beta I(L_x[X(s)] - 1)]^{-r} - [I + \beta I]^{-r}\right\}\left\{I - (I + \beta I)^{-r}\right\}^{-1}\mathbf{1}^T \end{aligned} \tag{3.156}$$

The Laplace transforms and probability generating functions are as shown in Subsection (3.3.2)

3.6.2 Compound phase type distributions probability generating functions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (3.152) the pgf of the compound distributions of PH-Zero-truncated Poisson with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH Zero truncated Poisson-Weibull	$Z(s) = \Upsilon \frac{\Lambda e^{\Lambda \frac{\beta}{\alpha} \frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}}} - I}{e^{\Lambda} - I} 1^T$
PH Zero truncated Poisson-Pareto	$Z(s) = \Upsilon \frac{\Lambda e^{\Lambda \sum_{k=0}^{\infty}(-1)^k \frac{\alpha}{\Gamma\alpha} \frac{\Gamma(\alpha+k)}{\beta^{2k+2}}} - I}{e^{\Lambda} - I} 1^T$
PH Zero truncated Poisson-Gen Pareto	$Z(s) = \Upsilon \frac{\Lambda e^{\Lambda \frac{1}{\lambda^{\gamma}\beta(\alpha,\gamma)} \sum_{k=0}^{\infty} \frac{-(\alpha+\gamma)}{\lambda^k} \frac{\Gamma\alpha+k}{s^{\alpha+k}}} - I}{e^{\Lambda} - I} 1^T$
PH Zero truncated Poisson-OPPL	$Z(s) = \Upsilon \frac{\Lambda e^{\Lambda \frac{\theta^2}{1+\theta} \frac{\theta+2-s}{[\theta+1-s]^2}} - I}{e^{\Lambda} - I} 1^T$
PH Zero truncated Poisson-TPPL	$Z(s) = \Upsilon \frac{\Lambda e^{\Lambda \frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}} - I}{e^{\Lambda} - I} 1^T$

Table 3.5. Compound phase type Zero truncated Poisson distributions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (3.154) the pgf of the compound distributions of PH-Zero truncated Geometric with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH ZT Geometric-Weibull	$Z(s) = \Upsilon P \; Q - Q^2 \frac{\beta}{\alpha} \frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}} \; \vec{1}^T$
PH ZT Geometric-Pareto	$Z(s) = \Upsilon P \; Q - Q^2 \sum_{k=0}^{\infty} (-1)^k \frac{\alpha}{\Gamma\alpha} \frac{\Gamma(\alpha+k)}{\beta^{2k+2}} \; \vec{1}^T$
PH ZT Geometric-Gen Pareto	$Z(s) = \Upsilon P \; Q - Q^2 \frac{1}{\lambda^{\gamma}\beta(\alpha,\gamma)} \sum_{k=0} \frac{}{\lambda^k} \frac{}{s^{\alpha+k}} \; \vec{1}^T$
PH ZT Geometric-OPPL	$Z(s) = \Upsilon P \; Q - Q^2 \frac{\theta^2}{1+\theta} \frac{}{[\theta+1-s]^2} \; \vec{1}^T$
PH ZT Geometric-TPPL	$Z(s) = \Upsilon P \; Q - Q^2 \frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2} \; 1^T$

Table 3.6. Compound phase type Zero truncated Geometric distributions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (3.156) the pgf of the compound distributions of PH-Zero-truncated Negative Binomial with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH Negative Binomial -Weibull	$Z(s) = \Upsilon \left[\left[I - \beta I \left(\frac{\beta}{\alpha} \frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta}-1]^{\beta}} \right)^{-1} - I \right]^{-r} - [I+\beta I]^{-r} \right] \left[I - (I+\beta I)^{-r} \right] \vec{1}^T$
PH Negative Binomial -Pareto	$Z(s) = \Upsilon \left[\left[I - \beta I \left(\sum_{k=0}^{\infty} (-1)^k \frac{\alpha\, \Gamma(\alpha+k)}{\Gamma\alpha\, \beta^{2k+2}} \right)^{-1} - I \right]^{-r} - [I+\beta I]^{-r} \right] \left[I - (I+\beta I)^{-r} \right] \vec{1}^T$
PH Negative Binomial -Gen Pareto	$Z(s) = \Upsilon \left[\left[I - \beta I \left(\frac{1}{\lambda^{\gamma}\beta(\alpha,\gamma)} \sum_{k=0}^{\infty} \frac{-(\alpha+\gamma)\,\Gamma\alpha+k}{\lambda^{k}-1\; s^{\alpha+k}} \right)^{-1} - I \right]^{-r} - [I+\beta I]^{-r} \right] \left[I - (I+\beta I)^{-r} \right] \vec{1}^T$
PH Negative Binomial -OPPL	$Z(s) = \Upsilon \left[\left[I - \beta I \left(\frac{\theta^2}{1+\theta} \frac{\theta+2-s}{[\theta+1-s]^2} \right)^{-1} - I \right]^{-r} - [I+\beta I]^{-r} \right] \left[I - (I+\beta I)^{-r} \right] \vec{1}^T$
PH Negative Binomial -TPPL	$Z(s) = \Upsilon \left[\left[I - \beta I \left(\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2} \right)^{-1} - I \right]^{-r} - [I+\beta I]^{-r} \right] \left[I - (I+\beta I)^{-r} \right] \vec{1}^T$

Table 3.7. Compound phase type Zero truncated Negative Binomial distributions

3.7 Phase type Panjer recursion formula for class $(a, b, 1)$

The pgf of the phase type distributions of aggregate losses,frequency distribution and severity distributions are used to derive phase type Panjer recursion formula. Let $Z(s)$, $Y(s)$ and $X(s)$ be $M * M$ matrix representing the phase type aggregate loss, PH count distribution pgf and phase type severity distribution pgf respectively.

Theorem 3.7.1 (Recursive form of phase type class $(a,\ b,\ 1)$ compound distribution). *If the distribution of N belongs to class $(a, b, 1)$ group then it follows the following recursive formula:*

$$Z(j) = \left[P_1 - (A+B)P_0 \right] y_j + \sum_{x=1}^{\infty} y(x) Z(j-x) \left(A + B\frac{x}{j} \right) \left[1 - Ay(0) \right]^{-1}$$

where $Z(j)$ is a $M * M$ matrix and $Z(0)$ is considered as the initial matrix.

Proof of theorem 3.7.1

The pgf of phase type aggregate loss distribution is expressed as:

$$Z(s) = E[s^{SN}] = \sum_{j=0}^{\infty} Z(j) s^j$$

The pgf of PH frequency distribution is expressed as:

$$Y(s) = E[s^N] = \sum_{n=0}^{\infty} P_n s^n$$

The pgf of phase type severity distribution is expressed as:

$$X(s) = E[s^{X_i}] \qquad = \sum_{x=0}^{\infty} y(x)s^x$$

Phase type aggregate loss distribution can be represented as a convolution of the pgf of phase type claim count distribution and pgf of phase type claim amount distribution as:

$$Z(s) = Y[X(s)] \tag{3.157}$$

The first derivative of equation (3.157)is:

$$Z'(s) = \{Y'[X'(s)]\}X'(s) \tag{3.158}$$

It is known that the pgf of PH frequency distribution is:

$$Y(s) = \sum_{n=0}^{\infty} P_n s^n \tag{3.159}$$

Combining equation (3.157) and equation (3.159) results to:

$$Z(s) = \sum_{n=0}^{\infty} P_n [X(s)]^n \tag{3.160}$$

The first derivative of the pgf of the phase type aggregate loss distribution as expressed in equation (3.160) is:

$$Z'(s) = \sum_{n=0}^{\infty} nP_n [X(s)]^{n-1} X'(s) \tag{3.161}$$

Phase type Panjer's model is expressed as :

$$P_n = \left(A + \frac{B}{n}\right) P_{n-1} \qquad n = 2, 3, 4, \ldots \tag{3.162}$$

Multiplying all through by $(n-1)$ equation (3.162) becomes:

$$(n-1)P_n = \left[(n-1)A + \frac{B}{n}(n-1)\right] P_{n-1}$$

$$nP_n - P_n = \left[An - \frac{B}{n} - A + B\right] P_{n-1}$$

$$nP_n - P_n = (An + B)P_{n-1} - \left(A + \frac{B}{n}\right) P_{n-1}$$

$$nP_n = (An + B)P_{n-1} + P_n - \left(A + \frac{B}{n}\right) P_{n-1} \tag{3.163}$$

Multiply equation (3.163) by $[X(s)]^{n-2}X'(s)$ and sum the outcome over n resulting to:

$$\sum_{n=2}^{\infty} nP_n[X(s)]^{n-2}X'(s) = \sum_{n=2}^{\infty}(nA + B)P_{n-1}[X(s)]^{n-2}X'(s) + \sum_{n=2}^{\infty}\left(P_n - \left(A + \frac{B}{n}\right)P_{n-1}\right)[X(s)]^{n-2}X'(s) \tag{3.164}$$

Combining equation (3.161) and equation (3.164) it becomes :

$$Z'(s) = \sum_{n=1}^{\infty}(n-1)AP_{n-1}[X(s)]^{n-1}X'(s) + (A+B)\sum_{n=1}^{\infty}P_{n-1}[X(s)]^{n-1}X'(s) + \sum_{n=1}^{\infty}P_n - \left(A + \frac{B}{n}\right)P_{n-1}[X(s)]^{n-1}X'(s) \tag{3.165}$$

By definition the pgf of phase type aggregate loss distribution is:

$$Z(s) = \sum_{j=0}^{\infty} X(j)s^j \tag{3.166}$$

The first derivative of equation (3.166) is :

$$Z'(s) = \sum_{j=1}^{\infty} jZ(j)s^{j-1} \tag{3.167}$$

By definition the pgf of phase type severity distribution is:

$$X(s) = \sum_{x=0}^{\infty} y(x)s^x \tag{3.168}$$

Equation's (3.168) first derivative is :

$$X'(s) = \sum_{x=1}^{\infty} xy(x)s^{x-1} \tag{3.169}$$

Merging equation (3.165) and equation (3.167) it becomes:

$$\sum_{j=1}^{\infty} jZ(j)s^{j-1} = A\{\sum_{n=1}^{\infty}(n-1)P_{n-1}[X(s)]^{n-1}\}X'(s) + (A+B)\{\sum_{n=1}^{\infty}P_{n-1}[X(s)]^{n-1}\}X'(s) + \sum_{n=1}^{\infty}P_n - \left(A + \frac{B}{n}\right)P_{n-1}[X(s)]^{n-1}X'(s) \tag{3.170}$$

Apart from equation (3.166) $Z(s)$ is also wriNen as:

$$Z(s) = \sum_{n=1}^{\infty} P_{n-1}[X(s)]^{n-1} \tag{3.171}$$

Hence equation's (3.171) derivative is ;

$$Z'(s) = \sum_{n=2}^{\infty}(n-1)P_{n-1}[X(s)]^{n-2}X'(s) \quad \textbf{(3.172)}$$

Combining equation (3.170), (3.171) and equation (3.172) results to:

$$\begin{aligned}\sum_{j=1}^{\infty} jZ(j)s^{j-1} &= AX(s)Z'(s) + (A+B)Z(s)X'(s) + \sum_{n=1}^{\infty}\left[P_n - \left(A+\frac{B}{n}\right)P_{n-1}\right][X(s)]^{n-1}X'(s)\\ &= A\sum_{x=0}^{\infty} y(x)s^x \sum_{j=1}^{\infty} jZ(j)s^{j-1} + (A+B)\sum_{j=0}^{\infty} Z(j)s^j \sum_{x=1}^{\infty} xy(x)s^{x-1}\\ &+ \sum_{n=1}^{\infty}\left[P_n - \left(A+\frac{B}{n}\right)P_{n-1}\right][X(s)]^{n-1}\sum_{x=1}^{\infty} xy(x)s^{x-1}\end{aligned} \quad \textbf{(3.173)}$$

Factoring out s from equation (3.173) results to :

$$\begin{aligned}\sum_{j=1}^{\infty} jZ(j)s^j &= A\sum_{x=0}^{\infty} y(x)s^x \sum_{j=1}^{\infty} jZ(j)s^j + (A+B)\sum_{j=0}^{\infty} Z(j)s^j \sum_{x=1}^{\infty} xy(x)s^x\\ &+ \sum_{n=1}^{\infty}\left[P_n - \left(A+\frac{B}{n}\right)P_{n-1}\right][X(s)]^{n-1}\sum_{x=1}^{\infty} xy(x)s^x\end{aligned} \quad \textbf{(3.174)}$$

But we know:

$$\sum_{j=1}^{\infty} jZ(j)s^j = \sum_{j=x+1}^{\infty} (j-x)Z(j-x)s^{j-x} \quad \textbf{(3.175)}$$

and

$$\sum_{j=0}^{\infty} Z(j)s_j = \sum_{j=x}^{\infty} Z(j-x)s^{j-x} \quad \textbf{(3.176)}$$

Combining equation (3.174), (3.175) and (3.176) results to:

$$\begin{aligned}\sum_{j=1}^{\infty} jZ(j)s^j &= A\sum_{x=0}^{\infty}\sum_{j=x+1}^{\infty} (j-x)y(x)Z(j-x)s^j + (A+B)\sum_{j=x}^{\infty}\sum_{x=1}^{\infty} Z(j-x)s^j xy(x)\\ &+ \sum_{n=1}^{\infty}\left[P_n - \left(A+\frac{B}{n}\right)P_{n-1}\right][X(s)]^{n-1}\sum_{x=1}^{\infty} xy(x)s^x\end{aligned} \quad \textbf{(3.177)}$$

Replacing equation (3.177) with $n = 1$ results to:

$$\begin{aligned} &= A\sum_{x=0}^{\infty}\sum_{j=x+1}^{\infty} (j-x)y(x)Z(j-x)s^j + (A+B)\sum_{j=x}^{\infty}\sum_{x=1}^{\infty} xy(x)Z(j-x)s^j + \left[P_1 - (A+B)P_0\right]y_j\\ &= A\sum_{x=0}^{\infty}\sum_{j=x+1}^{\infty} (j-x)y(x)Z(j-x)s^j + (A+B)\sum_{j=1}^{\infty}\sum_{x=1}^{\infty} xy(x)Z(j-x)s^j + \left[P_1 - (A+B)P_0\right]y_j\\ &= \sum_{j=1}^{\infty}\{A\sum_{x=0}^{\infty} (j-x)y(x)Z(j-x)\}s^j + \sum_{j=1}^{\infty}\{(A+B)\sum_{x=1}^{\infty} xy(x)Z(j-x)\}s^j + \left[P_1 - (A+B)P_0\right]y_j\end{aligned} \quad \textbf{(3.178)}$$

Combining the coefficients of s^j in equation (3.178) results to :

$$jZ(j) = A\sum_{x=0}^{\infty}(j-x)y(x)Z(j-x) + (A+B)\sum_{x=1}^{\infty}xy(x)Z(j-x) + \left[P_1-(A+B)P_0\right]y_j$$

$$= Ajy(0)Z(j) + A\sum_{x=1}^{\infty}(j-x)y(x)Z(j-x) + (A+B)\sum_{x=1}^{\infty}xy(x)Z(j-x)$$

$$+ \left[P_1-(A+B)P_0\right]y_j \tag{3.179}$$

Factoring out $I-Ay(0)$ in equation (3.179) results to:

$$\left[I-Ay(0)\right]jZ(j) = \sum_{x=1}^{\infty}\left[A(j-x)+(A+B)x\right]y(x)Z(j-x) + \left[P_1-(A+B)P_0\right]y_j$$

$$= \sum_{x=1}^{\infty}\left[Aj-Ax+Ax+Bx\right]y(x)Z(j-x) + \left[P_1-(A+B)P_0\right]y_j$$

$$\left[I-Ay(0)\right]jZ(j) = \sum_{x=1}^{\infty}\left(Aj+Bx\right)y(x)Z(j-x) + \left[P_1-(A+B)P_0\right]y_j \tag{3.180}$$

The recursive form of the compound distribution is found by dividing equation (3.180) by $\left[I-Ay(0)\right]j$ to get:

$$Z(j) = \left\{\left[P_1-(A+B)P_0\right]y_j + \sum_{x=1}^{\infty}\left(A+B\frac{x}{j}\right)y(x)Z(j-x)\right\}\left[I-Ay(0)\right]^{-1} \tag{3.181}$$

$\bar{Z}(j)$ can be expressed as a row vector as $\Upsilon Z(j)$ hence equation (3.181) becomes:

$$\bar{Z}(j) = \left\{\left[P_1-(A+B)P_0\right]y_j + \sum_{x=1}^{\infty}\left(A+B\frac{x}{j}\right)y(x)Z(j-x)\right\}\left[I-Ay(0)\right]^{-1} \tag{3.182}$$

The initial aggregate loss probability matrix is expressed as:

$$Z(0) = \sum_{n=0}^{\infty}\left[y(0)\right]^n P_n \tag{3.183}$$

$Z(j)$ can be expressed as a probability as:

$$z(j) = \Upsilon Z(j)\mathbf{1}^T$$

3.8 Chapter summary

The main objective of this chapter was to develop phase type Panjer class $(a, b, 0)$ distributions, phase type Panjer class $(a, b, 1)$ distributions and their compound phase type probability generating functions. Phase type distributions of Panjer class $(a, b, 0)$ and class $(a, b, 1)$ are developed using pgf technique as well as their compound phase type probability generating functions. Panjer recursive formulas for both classes are also developed.

4 PHASE TYPE POISSON LINDLEY AND ZERO-TRUNCATED POISSON LINDLEY DISTRIBUTIONS

4.1 Introduction

In this chapter phase type one parameter and two parameter Poisson lindley distributions can be derived considering a mixing distribution which follows phase type lindley distribution. Phase type Zero-truncated one parameter and two parameter Poisson Lindley distributions are also derived when the prior distribution follows size-biased Poisson distribution (SBPD). These phase type distributions are used to estimate claim frequency probabilities of secondary cancer cases. These distributions are phase type distribution hence they in cooperates matrices. The matrix inverse of these matrix parameters have been wriNen as fractions for simplification of the distributions appearance and understanding. Matrices of these phase type distributions are derived using Chapman-Kolmogorov equations as multi-state models hence in-cooperating secondary cancer transitions in claim frequency. This improves projection of aggregate losses as it also in-cooperates the dynamic nature of cancer and also heterogeneity aspect of claim data as they are mixture distributions. The phase type Zero-truncated Poisson Lindley further affect estimation of claim count data as the do not in-cooperate zero claim counts which can not aNract any claim severity amounts.

4.2 Phase type one parameter Poisson Lindley distribution

Definition 4.2.1. *A random variable X is considered a PH one parameter Poisson Lindley distribution if:*
$X|\gamma \sim Po(\gamma) \qquad \gamma|\Gamma \sim PH - OPL(\Gamma)$
for $\gamma > 0$ *and* Γ *is* $M * M$ *matrix.*

Theorem 4.2.2. *If* $X \sim PH - OPPL$ *distribution the pdf of* X *can be given as:*

$$f(x;\Gamma) = \vec{\Upsilon}\frac{\Gamma^2}{(I+\Gamma)^{x+3}}\{(x+2)I+\Gamma\}\vec{1}^{T} \tag{4.1}$$

where Γ *represents* $M * M$ *and* I *represents an identity matrix.*

Proof of theorem 4.2.2

If $X|\gamma \sim Po(\gamma)$ and $\gamma|\Gamma \sim PH - OPL(\Gamma)$, then the pdf of the variable X is illustrated by;

$$P(x) = \int_0^{\infty} Pr(x|\gamma)\, f(\gamma;\Gamma)d\gamma$$

where $f(\gamma;\Gamma)$ is $PH-OPL(\Gamma)$.

$$P(x)=\int_0^\infty \frac{e^{-\gamma}\gamma^x}{x!}\frac{\Gamma^2}{I+\Gamma}(1+\gamma)e^{-\Gamma\gamma}\,d\gamma \qquad \gamma>0, \Gamma=M*M$$

$$P(x)=\frac{\Gamma^2}{I+\Gamma}\int_0^\infty\left[\frac{\gamma^x}{x!}e^{-\gamma(I+\Gamma)}+\frac{\gamma^{x+1}}{x!}e^{-\gamma(I+\Gamma)}\right]d\gamma \qquad =\Upsilon\frac{\Gamma^2}{(I+\Gamma)^{x+3}}\{(x+2)I+\Gamma\}\mathbf{1}^T \tag{4.2}$$

4.2.1 Properties of phase type one parameter Poisson Lindley distribution

The r^{th} moments of PH-OPPL distribution expressed as:

$$E(X^r)=\int_0^\infty x^r f(x,\Gamma)dx \qquad =\frac{\Gamma^2}{I+\Gamma}\int_0^\infty x^r e^{-\Gamma x}(1+x)dx \qquad =\Upsilon\frac{x![(x+1)I+\Gamma]}{(\Gamma+I)}\mathbf{1}^T \tag{4.3}$$

The mean and variance of PH-OPPL distribution is derived from equation (4.3) as:

(i) Expectation

$$E(x)=\frac{1![(1+1)I+\Gamma]}{\Gamma(\Gamma+I)} \qquad =\Upsilon\frac{(2I+\Gamma)}{\Gamma(\Gamma+I)}\mathbf{1}^T \tag{4.4}$$

(ii) Variance

$$Var(x)=\frac{2![(2+1)I+\Gamma]}{\Gamma^2(\Gamma+I)}-\left(\frac{(2I+\Gamma)}{\Gamma(\Gamma+I)}\right)^2 \qquad =\Upsilon\left(\frac{2I+4\Gamma+\Gamma^2}{(\Gamma+I)^2}+\frac{2I+\Gamma}{\Gamma(\Gamma+I)}\right)\mathbf{1}^T \tag{4.5}$$

The probability generating function of PH-OPPL distribution is expressed as:

$$X(s)=\int_0^\infty e^{\gamma(1-s)}\frac{\Gamma^2}{I+\Gamma}(1+\gamma)e^{-\Gamma\gamma}d\gamma \qquad =\frac{\Gamma^2}{I+\Gamma}\left[\int_0^\infty \gamma e^{-\gamma(I+\Gamma-sI)}d\gamma+\int_0^\infty e^{-\gamma(\Gamma+I-sI)}d\gamma\right]$$

$$=\Upsilon\frac{\Gamma^2}{I+\Gamma}\left[\frac{\Gamma+(2-s)I}{[\Gamma+(1-s)I]^2}\right]\mathbf{1}^T \tag{4.6}$$

4.3 Phase type two parameter Poisson Lindley distribution

Definition 4.3.1. *A random variable X is considered a phase type two parameter Poisson Lindley distribution if it satisfies:*

$X|\gamma \sim Po(\gamma) \qquad \gamma|\Gamma,\alpha \sim PH-TPL(\Gamma,\alpha)$

for $\alpha>0,\gamma>0$ and Γ represents $M*M$ matrix.

Theorem 4.3.2. *If $X\sim PH-TPPL$ distribution, the probability density function of X is given by:*

$$f(x;\Gamma,\alpha)=\Upsilon\frac{\Gamma^2}{(I+\Gamma)^{x+2}}\left[I+\frac{(\alpha+x)I}{\alpha\Gamma+I}\right]\mathbf{1}^T \tag{4.7}$$

where $\alpha>0,\Gamma$ is $M*M$ and I is an identity matrix.

Proof theorem 4.3.2

If $X|\gamma \sim Po(\gamma)$ and $\gamma|\Gamma,\alpha \sim PH-TPL(\Gamma,\alpha)$, then the pdf of variable X is expressed as;

$$P(x) = \int_0^\infty Pr(X = x|\gamma) f(\gamma;\Gamma,\alpha) d\gamma$$

where $f(\gamma;\Gamma,\alpha)$ is $PH-TPL(\Gamma,\alpha)$.

$$P(x) = \int_0^\infty \frac{e^{-\gamma}\gamma^x}{x!}\frac{\Gamma^2}{I+\Gamma\alpha}(\alpha+\gamma)e^{-\Gamma\gamma}d\gamma \qquad \gamma > 0,\ \Gamma = M * M$$

$$P(x) = \frac{\Gamma^2}{\alpha\Gamma+I}\left[\int_0^\infty \frac{\gamma^x}{x!}\alpha e^{-\gamma(I+\Gamma)}\,d\gamma + \int_0^\infty \frac{\gamma^{x+1}}{x!}e^{-\gamma(I+\Gamma)}\,d\gamma\right] = \Upsilon\frac{\Gamma^2}{(I+\Gamma)^{x+2}}\left[I+\frac{(\alpha+x)I}{(\alpha\Gamma+I)}\right]1^T \qquad \textbf{(4.8)}$$

4.3.1 Properties of phase type two parameter Poisson Lindley distribution

The r^{th} moments of PH-TPPL distribution is illustrated as:

$$E(X^r) = \int_0^\infty x^r f(x,\Gamma,\alpha)dx = \int_0^\infty\left[\sum_{x=0}^\infty x^r\frac{e^{-\gamma}\gamma^x}{x!}\right]\frac{\Gamma^2}{I+\alpha\Gamma}(\alpha+\gamma)e^{-\Gamma\gamma}\,d\gamma$$

$$P(x) = \frac{\Gamma^2}{I+\Gamma\alpha}\left[\alpha\frac{\Gamma(r+1)}{\Gamma^{r+1}} + \frac{\Gamma(r+2)}{\Gamma^{r+2}}\right] = \Upsilon\frac{\Gamma(r+1)I}{\Gamma^r}\frac{\alpha\Gamma+(r+1)I}{\alpha\Gamma+I}1^T \qquad \textbf{(4.9)}$$

The mean and variance of PH-TPPL distribution is derived from equation (4.9) as:

(i) Expectation

$$E(x) = \frac{\Gamma^2}{\alpha\Gamma+I}\int_0^\infty \gamma(\alpha+\gamma)e^{-\Gamma\gamma}\,d\gamma = \Upsilon\frac{(2I+\Gamma\alpha)}{\Gamma(\Gamma\alpha+I)}1^T \qquad \textbf{(4.10)}$$

(ii) Variance

$$Var(x) = E(x^2) - [E(x)]^2$$

$$E(x^2) = \frac{\alpha\Gamma+2I}{\Gamma(\alpha\Gamma+I)} + \frac{2(\alpha\Gamma+3I)}{\Gamma^2(\alpha\Gamma+I)} = \Upsilon\left[\frac{\alpha\Gamma+2I}{\Gamma(\alpha\Gamma+I)} + \frac{2(\alpha\Gamma+3I)}{\Gamma^2(\alpha\Gamma+I)} - \left(\frac{(2I+\Gamma\alpha)}{\Gamma(\Gamma\alpha+I)}\right)^2\right]1^T \qquad \textbf{(4.11)}$$

The probability generating function of PH-TPPL distribution can be expressed as:

$$X(s) = \frac{\Gamma^2}{(\Gamma+I)^2}\sum_{x=0}^\infty\left[\frac{sI}{(\Gamma+I)}\right]^x + \frac{\Gamma^2}{(\Gamma+I)^2(\alpha\Gamma+I)}\sum_0^\infty(\alpha+x)\left[\frac{s}{\Gamma+I}\right]^x$$

$$= \Upsilon\frac{\alpha\Gamma[\Gamma+(1-s)I]+\Gamma^2}{(\alpha\Gamma+I)[\Gamma+(1-s)I]^2}1^T \qquad \textbf{(4.12)}$$

The value of Γ has been calculated hence the value of α can be estimated from equation (4.10) having calculated the value of $E(x)$.

4.4 Compound distributions of one parameter and two parameter Poisson Lindley distributions with severity distributions

Compound phase type distributions (CPHD) considered in modeling secondary cancer cases using one parameter and two parameter Poisson Lindley distributions are developed in this section. These distributions are mixture distributions hence they factor in the heterogeneity aspect of claim count data. Compound distribution's probability generating functions are derived by convoluting probability generating function of claim count and claim amount distributions as expressed in definition (3.3.1).

4.4.1 General expression of phase type compound distributions

Theorem 4.4.1 (Compound phase type one parameter Poisson Lindley distribution)**.** *If the pgf of $N \sim PH - OPPL(\Gamma)$ the compound pgf of N is:*

$$Z(S) = \vec{Y}\,\frac{\Gamma^2}{I+\Gamma}\left[\frac{\Gamma + (2 - L_x[X(S)])I}{[\Gamma + (1 - L_x[X(S)]]I]^2}\right]\vec{1}^{T} \tag{4.13}$$

where $L_x[X(S)]$ is as defined in theorem (3.3.2).

Proof of theorem 4.4.1

Let the pgf of the compound distribution be expressed as $Z(s) = F[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (4.6) it becomes:

$$\begin{aligned} Z(S) &= F[L_x[X(S)]] \\ &= \vec{Y}\,\frac{\Gamma^2}{I+\Gamma}\left[\frac{\Gamma + (2 - L_x[X(S)])I}{[\Gamma + (1 - L_x[X(S)]]I]^2}\right]\vec{1}^{T} \end{aligned} \tag{4.14}$$

Theorem 4.4.2 (Compound phase type two parameter Poisson Lindley distribution)**.** *If the pgf of $N \sim PH - TPPL(\Gamma)$ the the compound pgf of N is:*

$$Z(S) = \vec{Y}\,\frac{\alpha\Gamma[\Gamma + (1 - L_x[X(S)])I] + \Gamma^2}{(\alpha\Gamma + I)[\Gamma + (1 - L_x[X(S)])I]^2}\vec{1}^{T} \tag{4.15}$$

where $L_x[X(S)]$ is as defined in theorem (3.3.2).

Proof of theorem 4.4.2

Let the pgf of the compound distribution be expressed as $Z(s) = F[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (4.12) it becomes:

$$\begin{aligned} Z(S) &= Y[L_x[X(S)]] \\ &= \vec{Y}\,\frac{\alpha\Gamma[\Gamma + (1 - L_x[X(S)])I] + \Gamma^2}{(\alpha\Gamma + I)[\Gamma + (1 - L_x[X(S)])I]^2}\vec{1}^{T} \end{aligned} \tag{4.16}$$

4.4.2 Compound phase type distributions probability generating functions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (4.14) the pgf of the compound distributions of PH-one parameter Poisson Lindley with severity distributions:

Distributions	Pgf of compound phase type distributions
PH OPPL-Weibull	$Z(S)=\Upsilon\left[\frac{\Gamma^2}{I+\Gamma}\frac{\Gamma+(2-\frac{\beta}{\alpha}\frac{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]\beta}{\Gamma\beta})I}{[\Gamma+(1-\frac{\beta}{\alpha}\frac{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]\beta}{\Gamma\beta})I]^2}\right]^{-1}\mathbf{1}^T$
PH OPPL-Pareto	$Z(S)=\Upsilon\left[\frac{\Gamma^2}{I+\Gamma}\frac{\Gamma+(2-\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}})I}{[\Gamma+(1-\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}})I]^2}\right]^{-1}\mathbf{1}^T$
PH OPPL-Gen Pareto	$Z(S)=\Upsilon\left[\frac{\Gamma^2}{I+\Gamma}\frac{\Gamma+(2-\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}})I}{[\Gamma+(1-\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}})I]^2}\right]^{-1}\mathbf{1}^T$
PH OPPL-OPPL	$Z(S)=\Upsilon\left[\frac{\Gamma^2}{I+\Gamma}\frac{\Gamma+\left(2-\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}\right)I}{\left[\Gamma+\left(1-\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}\right)I\right]^2}\right]^{-1}\mathbf{1}^T$
PH OPPL-TPPL	$Z(S)=\Upsilon\left[\frac{\Gamma^2}{I+\Gamma}\frac{\Gamma+\left(2-\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}\right)I}{\left[\Gamma+\left(1-\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}\right)I\right]^2}\right]^{-1}\mathbf{1}^T$

Table 4.1. Compound phase type one parameter Poisson Lindley distributions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (4.16) the pgf of the compound distributions of PH-two parameter Poisson Lindley with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH TPPL-Weibull	$Z(S)=\Upsilon\left[\frac{\alpha\Gamma\left[\Gamma+\left(1-\frac{\beta}{\alpha}\frac{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]\beta}{\Gamma\beta}\right)I\right]+\Gamma^2}{(\alpha\Gamma+I)\left[\Gamma+\left(1-\frac{\beta}{\alpha}\frac{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]\beta}{\Gamma\beta}\right)I\right]^2}\right]^{-1}\mathbf{1}^T$
PH TPPL-Pareto	$Z(S)=\Upsilon\left[\frac{\alpha\Gamma\left[\Gamma+\left(1-\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}\right)I\right]+\Gamma^2}{(\alpha\Gamma+I)\left[\Gamma+\left(1-\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}\right)I\right]^2}\right]^{-1}\mathbf{1}^T$
PH TPPL-Gen Pareto	$Z(S)=\Upsilon\left[\frac{\alpha\Gamma\left[\Gamma+\left(1-\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}}\right)I\right]+\Gamma^2}{(\alpha\Gamma+I)\left[\Gamma+\left(1-\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}}\right)I\right]^2}\right]^{-1}\mathbf{1}^T$
PH TPPL-OPPL	$Z(S)=\Upsilon\left[\frac{\alpha\Gamma\left[\Gamma+\left(1-\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}\right)I\right]+\Gamma^2}{(\alpha\Gamma+I)\left[\Gamma+\left(1-\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}\right)I\right]^2}\right]^{-1}\mathbf{1}^T$
PH TPPL-TPPL	$Z(S)=\Upsilon\left[\frac{\alpha\Gamma\left[\Gamma+\left(1-\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}\right)I\right]+\Gamma^2}{(\alpha\Gamma+I)\left[\Gamma+\left(1-\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}\right)I\right]^2}\right]^{-1}\mathbf{1}^T$

Table 4.2. Compound phase type two parameter Poisson Lindley distributions

4.5 Phase type Zero-truncated one parameter Poisson Lindley distribution

Definition 4.5.1. *A random variable X is considered a phase type Zero-truncated one parameter Poisson Lindley distribution if $x|\gamma$ follows size-biased Poisson distribution illustrated as:*

$$f(x|\gamma) = \frac{e^{-\gamma}\gamma^{x-1}}{(x-1)!} \tag{4.17}$$

for $\gamma > 0$
where parameter γ satisfies the expression ;

$$g(\gamma;\Gamma) = \frac{\Gamma^2}{\Gamma^2+3\Gamma+I}\left[(\Gamma+I)\gamma+(\Gamma+2I)\right]e^{-\Gamma\gamma} \tag{4.18}$$

for $\gamma > 0$, Γ is $M*M$ matrix and I is an identity matrix.

Theorem 4.5.2. *If $X \sim PH-ZTOPPL$ distribution then the probability mass function of X is:*

$$f(x;\Gamma) = \vec{\Upsilon}\frac{\Gamma^2}{\Gamma^2+3\Gamma+I}\left(\frac{(x+2)I+\Gamma}{(\Gamma+I)^x}\right)\vec{1}^T \qquad x = 1,2,3,.... \tag{4.19}$$

where Γ is $M*M$ and I is an identity matrix.

Proof of theorem 4.5.2

Given the prior distribution is $X|\gamma \sim SBPD(\gamma)$, then the pmf of variable X of PH-ZTOPPL distribution is illustrated as:

$$\begin{aligned}
P(x) &= \int_0^\infty f(x|\gamma)g(\gamma;\Gamma)d\gamma = \int_0^\infty \frac{e^{-\gamma}\gamma^{x-1}}{(x-1)!}\cdot\frac{\Gamma^2}{\Gamma^2+3\Gamma+I}\left[(\Gamma+I)\gamma+(\Gamma+2I)\right]e^{-\Gamma\gamma}d\gamma \\
&= \frac{\Gamma^2}{(\Gamma^2+3\Gamma+I)(x-1)!}\int_0^\infty e^{-\gamma(I+\Gamma)} * \left[(\Gamma+I)\gamma^{x}+(\Gamma+2I)\gamma^{x-1}\right]d\gamma \\
&= \vec{\Upsilon}\frac{\Gamma^2}{\Gamma^2+3\Gamma+I}\left(\frac{(x+2)I+\Gamma}{(\Gamma+I)^x}\right)\vec{1}^T \qquad x = 1,2,3,....
\end{aligned} \tag{4.20}$$

Properties of phase type Zero-truncated one parameter Poisson Lindley distribution

The r^{th} moments of PH-ZTOPPL distribution is given by:

$$\begin{aligned}
E(X^r) &= \int_0^\infty x^r f(x,\Gamma)d\gamma = \int_0^\infty\left[\sum_{x=1}^{\infty} x^r\frac{e^{-\gamma}\gamma^{x-1}}{(x-1)!}\right]\frac{\Gamma^2}{\Gamma^2+3\Gamma+I}\left[(\Gamma+I)\gamma+(\Gamma+2I)\right]e^{-\Gamma\gamma}d\gamma \\
&= \frac{\Gamma^2}{\Gamma^2+3\Gamma+I}\int_0^\infty \gamma^{r-1}(\gamma+r)[(\Gamma+I)\gamma+(\Gamma+2I)]e^{-\Gamma\gamma}d\gamma \\
&= \vec{\Upsilon}\frac{r!(\Gamma+I)^2[(r+1)I+\Gamma]}{\Gamma^r[\Gamma^2+3\Gamma+I]}\vec{1}^T
\end{aligned} \tag{4.21}$$

The mean and variance of PH-ZTOPPL distribution id derived obtained from equation (4.21) as:

(i) Expectation

$$E(x) = \frac{1!(\Gamma + I)^2[2I + \Gamma]}{\Gamma^1[\Gamma^2 + 3\Gamma + I]} = \vec{\Upsilon}\, \frac{(\Gamma + I)^2[2I + \Gamma]}{\Gamma[\Gamma^2 + 3\Gamma + I]} \vec{1}^T \tag{4.22}$$

(ii) Variance

$$Var(x) = E[x^2] - (E[x])^2 \quad E[x^2] = \frac{2!(\Gamma + I)^2[(3I + \Gamma]}{\Gamma^2[\Gamma^2 + 3\Gamma + I]} = \frac{(\Gamma + I)^2[\Gamma^2 + 4\Gamma + 6I]}{\Gamma^2[\Gamma^2 + 3\Gamma + I]}$$

$$Var(x) = \Upsilon \left[\frac{(\Gamma + I)^2[\Gamma^2 + 4\Gamma + 6I]}{\Gamma^2[\Gamma^2 + 3\Gamma + I]} - \left(\frac{(\Gamma + I)^2[2I + \Gamma]}{\Gamma[\Gamma^2 + 3\Gamma + I]} \right)^2 \right] \mathbf{1}^T \tag{4.23}$$

The probability generating function of PH-ZTOPPL distribution is expressed as:

$$X(s) = E[s^x] = \frac{\Gamma^2}{\Gamma^2 + 3\Gamma + I} \sum_{x=1}^{\infty} s^x \frac{xI + \Gamma + 2I}{(\Gamma + I)^x}$$

$$= \frac{\Gamma^2}{\Gamma^2 + 3\Gamma + I} \left[\sum_{x=1}^{\infty} x \left(\frac{sI}{\Gamma + I} \right)^x + (\Gamma + 2I) \sum_{x=1}^{\infty} \left(\frac{sI}{\Gamma + I} \right)^x \right]$$

$$= \vec{\Upsilon} \frac{\Gamma^2 s}{\Gamma^2 + 3\Gamma + I} \left[\frac{(\Gamma + I)}{(\Gamma + I - sI)^2} + \frac{(\Gamma + 2I)}{\Gamma + I - sI} \right] \vec{1}^T \tag{4.24}$$

4.6 Phase type Zero-truncated two parameter Poisson Lindley distribution

Definition 4.6.1. *A random variable X is considered a phase type Zero-truncated two parameter Poisson Lindley distribution if $x|\gamma$ follows size-biased Poisson distribution which can be represented as:*

$$f(x|\gamma) = \frac{e^{-\gamma}\gamma^{x-1}}{\Gamma(x)} \tag{4.25}$$

for $\gamma > 0$
where parameter γ satisfies the function ;

$$g(\gamma; \Gamma, \alpha) = \frac{\Gamma^2}{\Gamma^2\alpha + \Gamma\alpha + 2\Gamma + I} \left((\Gamma + I)\gamma + \alpha(\Gamma + I) + I \right) e^{-\Gamma\gamma} \tag{4.26}$$

for $\gamma > 0$, $\Gamma^2\alpha + \Gamma\alpha + 2\Gamma + I$ $\Gamma > 0$ represents $M * M$ matrix and I represents an identity matrix.

Theorem 4.6.2. *If $X \sim PH - ZTTPPL$ distribution then the probability mass function of X is expressed as:*

$$f(x; \Gamma, \alpha) = \vec{\Upsilon} \frac{\Gamma^2}{\Gamma^2\alpha + \Gamma\alpha + 2\Gamma + I} \frac{(x + 1)I + \alpha(\Gamma + I)}{(\Gamma + I)^x} \vec{1}^T \qquad x = 1, 2, 3, \ldots \tag{4.27}$$

where $\alpha > 0$,Γ represents $M * M$ and I represents an identity matrix.

Proof of theorem 4.6.2

Given the prior distribution as $X|\gamma \sim SBPD(\gamma)$, then the pmf of variable X of PH-ZTTPPL distribution is illustrated as:

$$P(x) = \int_0^\infty f(x|\gamma)g(\gamma;\Gamma)d\gamma$$

$$= \int_0^\infty \frac{e^{-\gamma}\gamma^{x-1}}{\Gamma(x)} * \frac{\Gamma^2}{\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I}\left((\Gamma+I)\gamma+\alpha(\Gamma+I)+I\right)e^{-\Gamma\gamma}d\gamma$$

$$= \frac{\Gamma^2}{[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]\Gamma(x)}\int_0^\infty\left[(\Gamma+I)\gamma^x+\left(\alpha(\Gamma+I)+I\right)\gamma^{x-1}\right]e^{-(\Gamma+I)\gamma}d\gamma$$

$$= \vec{\Upsilon}\frac{\Gamma^2}{[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]}\frac{(x+1)I+\alpha(\Gamma+I)}{(\Gamma+I)^x}\vec{1}^T \qquad x = 1, 2, 3, \ldots \qquad \textbf{(4.28)}$$

4.6.1 Properties of phase type Zero-truncated two parameter Poisson Lindley distribution

The r^{th} moments of PH-ZTTPPL distribution is given by:

$$E(X^r) = E[E(x^r|\gamma)] = \int_0^\infty\left[\sum_{x=1}^{\infty} x^r\frac{e^{-\gamma}\gamma^{x-1}}{(x-1)!}\right] * [(\gamma+I)\gamma+\alpha(\gamma+I)+I]e^{-\gamma\gamma}d\gamma$$

$$= \frac{\Gamma^2}{[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]}\int_0^\infty \gamma^{r-1}(\gamma+r)\left[(\gamma+I)\gamma+\alpha(\gamma+I)+I\right]e^{-\Gamma\gamma}d\gamma$$

$$= \vec{\Upsilon}\frac{r!(\Gamma+I)^2[\Gamma\alpha+(r+1)I]}{\Gamma^r[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]}\vec{1}^T \qquad 1, 2, 3, \ldots \qquad \textbf{(4.29)}$$

The mean and variance of PH-ZTTPPL distribution is derived from equation (4.29) as:

(i) Expectation

$$E(x) = \frac{1!(\Gamma+I)^2[\Gamma\alpha+2I]}{\Gamma^1[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]} = \vec{\Upsilon}\frac{(\Gamma+I)^2[\Gamma\alpha+2I]}{\Gamma[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]}\vec{1}^T \qquad \textbf{(4.30)}$$

(ii) Variance

$$Var(x) = E[x^2]-(E[x])^2 \qquad E[x^2] = \frac{2!(\Gamma+I)^2[\Gamma\alpha+3I]}{\Gamma^2[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]}$$

$$E[x^2] = \frac{2!(\Gamma+I)^2[\Gamma\alpha+3I]}{\Gamma^2[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]} - \left(\frac{(\Gamma+I)^2[\Gamma\alpha+2I]}{\Gamma[\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I]}\right)^2$$

$$Var(x) = \vec{\Upsilon}\frac{(\Gamma+I)^2\left[\Gamma^3\alpha^2+(\alpha+5)\Gamma^2\alpha+(4\alpha+6)\Gamma+2I\right]}{\Gamma^2[\Gamma^2\alpha+\Gamma\alpha2\Gamma+I]^2}\vec{1}^T \qquad \textbf{(4.31)}$$

The probability generating function of PH-ZTTPPL distribution is expressed as:

$$X(s) = E[s^x] = \sum_{x=1}^{\infty} s^x \left[\frac{\Gamma^2}{\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I}\,\frac{xI\alpha(\Gamma+I)+I}{(\Gamma+I)^x}\right]$$

$$= \vec{\Upsilon}\,\frac{\Gamma^2 s}{\Gamma^2\alpha+\Gamma\alpha+2\Gamma+I}\left[\frac{(\Gamma+I)}{(\Gamma+I-sI)^2}+\frac{[\alpha(\Gamma+I)+I]}{(\Gamma+I+sI)}\right]\vec{1}^{T} \tag{4.32}$$

The value of Γ has been calculated hence the value α is derived from equation (4.30) considering $E(x)$ is a known value.

4.7 Compound distributions of Zero-truncated one parameter and Zero-truncated two parameter Poisson Lindley distributions with severity distributions

Compound phase type distributions (CPHD) considered in modeling secondary cancer cases using Zero-truncated one parameter and two parameter Poisson Lindley distributions are derived in this section. These distributions are mixture distributions hence they factor in the heterogeneity aspect of claim count data. Zero truncated distributions do not take in to account zero claim count which is the case of real claim data as zero claim count can not aNract any claim severity amount. Compound distribution's probability generating functions are derived by convoluting of probability generating function of claim count and claim amount distributions as expressed in definition (3.3.1).

4.7.1 General expression of phase type compound distributions

Theorem 4.7.1 (Compound phase type Zero-truncated one parameter Poisson Lindley distribution). *If the pgf of $N \sim PH - ZTOPPL(\Gamma)$ the compound pgf of N is:*

$$Z(S) = \Upsilon\left[\frac{\Gamma^2 L_x[X(S)]}{\Gamma^2+3\Gamma+I}\left(\frac{(\Gamma+I)}{(\Gamma+I-L_x[X(S)]I)^2}+\frac{(\Gamma+2I)}{\Gamma+I-L_x[X(S)]I}\right)\right]1^{T} \tag{4.33}$$

where $L_x[X(s)]$ is as the Laplace transform.

Proof of theorem 4.7.1

Let the pgf of the compound distribution be expressed as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (4.24) it becomes:

$$Z(S) = Y[L_x[X(S)]]$$

$$= \vec{\Upsilon}\left[\frac{\Gamma^2 L_x[X(S)]}{\Gamma^2+3\Gamma+I}\left(\frac{(\Gamma+I)}{(\Gamma+I-L_x[X(S)]I)^2}+\frac{(\Gamma+2I)}{\Gamma+I-L_x[X(S)]I}\right)\right]\vec{1}^{T} \tag{4.34}$$

Theorem 4.7.2 (Compound phase type Zero-truncated two parameter Poisson Lindley distribution)**.** *If the pgf of* $N \sim PH - ZTTPPL(\Gamma, \alpha)$ *the the compound pgf of* N *is:*

$$Z(S) = Y\frac{\Gamma^2 L_x[X(S)]}{\Gamma^2\alpha + \Gamma\alpha + 2\Gamma + I}\left[\frac{(\Gamma + I)}{(\Gamma + I - L_x[X(S)]I)^2} + \frac{[\alpha(\Gamma + I) + I]}{(\Gamma + I - L_x[X(S)]I)}\right]\vec{1}^T \qquad \textit{(4.35)}$$

where $L_x[X(s)]$ is as the Laplace transform.

Proof of theorem 4.7.2

Let the pgf of the compound distribution be expressed as $Z(s) = Y[X(s)]$, hence replacing the pgf of frequency distribution as shown in equation (4.32) it becomes:

$$\begin{aligned} Z(S) &= Y[L_x[X(S)]] \\ &= \vec{Y}\frac{\Gamma^2 L_x[X(S)]}{\Gamma^2\alpha + \Gamma\alpha + 2\Gamma + I}\left[\frac{(\Gamma + I)}{(\Gamma + I - L_x[X(S)]I)^2} + \frac{[\alpha(\Gamma + I) + I]}{(\Gamma + I - L_x[X(S)]I)}\right]\vec{1}^T \end{aligned} \qquad \textbf{(4.36)}$$

4.7.2 Compound phase type distributions probability generating functions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (4.34) the pgf of the compound distributions of PH-one parameter Poisson Lindley with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH ZT OPPL - Weibull	$Z(S)=\Upsilon\frac{\Gamma^2\frac{\beta}{\alpha}\frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\frac{\beta}{\alpha}\frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\frac{\beta}{\alpha}\frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}}I}\right]^{-1}T$
PH ZT OPPL -Pareto	$Z(S)=\Upsilon\frac{\Gamma^2\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}I}\right]^{-1}T$
PH ZT OPPL -Gen Pareto	$Z(S)=\Upsilon\frac{\Gamma^2\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}}I}\right]^{-1}T$
PH ZT OPPL -OPPL	$Z(S)=\Upsilon\frac{\Gamma^2\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}I}\right]^{-1}T$
PH ZT OPPL -TPPL	$Z(S)=\Upsilon\frac{\Gamma^2\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}I}\right]^{-1}T$

Table 4.3. Compound phase type zero truncated one parameter Poisson Lindley distributions

Replacing equation (3.81), (3.82), (3.83), (3.84), and (3.85) in equation (4.36) the pgf of the compound distributions of PH-two parameter Poisson Lindley with severity distributions are:

Distributions	Pgf of compound phase type distributions
PH ZT TPPL -Weibull	$Z(S)=\Upsilon\frac{\Gamma^2\frac{\beta}{\alpha}\frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\frac{\beta}{\alpha}\frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\frac{\beta}{\alpha}\frac{\Gamma\beta}{[s\alpha+(\frac{x}{\alpha})^{\beta-1}]^{\beta}}I}\right]^{-1}T$
PH ZT TPPL -Pareto	$Z(S)=\Upsilon\frac{\Gamma^2\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\sum_{k=0}^{\infty}(-1)^k\frac{\alpha}{\Gamma\alpha}\frac{\Gamma(\alpha+k)}{\beta^{2k+2}}I}\right]^{-1}T$
PH ZT TPPL -Gen Pareto	$Z(S)=\Upsilon\frac{\Gamma^2\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\frac{1}{\gamma^{\gamma}\beta(\alpha,\gamma)}\sum_{k=0}^{\infty}\frac{-(\alpha+\gamma)}{\gamma^k}\frac{\Gamma\alpha+k}{s^{\alpha+k}}I}\right]^{-1}T$
PH ZT TPPL -OPPL	$Z(S)=\Upsilon\frac{\Gamma^2\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\frac{\theta^2}{1+\theta}\frac{\theta+2-s}{[\theta+1-s]^2}I}\right]^{-1}T$
PH ZT TPPL -TPPL	$Z(S)=\Upsilon\frac{\Gamma^2\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}}{\Gamma^2+3\Gamma+I}\left[\frac{(\Gamma+I)}{\left(\Gamma+I-\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}I\right)^2}+\frac{(\Gamma+2I)}{\Gamma+I-\frac{\alpha\theta[\theta+1-s]+\theta^2}{(\alpha\theta+1)[\theta+1-s]^2}I}\right]^{-1}T$

Table 4.4. Compound phase type zero truncated two parameter Poisson Lindley distributions

4.8 Chapter summary

The main objective of this chapter was to develop phase type Poisson Lindley distributions, Phase type zero truncated Poisson Lindley distributions and compound phase type probability generating functions. Phase type Poisson Lindley and Phase type zero truncated Poisson Lindley are constructed as well as their compound phase type probability generating functions.

5 MULTI-STATE CANCER MODEL

5.1 Introduction

In this chapter, we outline the mathematics of applying a Markov framework to multiple state models for secondary cancer insurance products. The Markov framework is used to determine the paNern of secondary cancer models and therefore the transition probabilities as well as the transition intensities are estimated. The transition matrix obtained improves on the estimation of aggregate losses of cancer as it factors in the movement between different states. Only Markov models with finite states are be considered in this research. Multi-state models derived in this chapter represents the matrix parameters of the phase type distributions of class $(a, b, 0)$ and class $(a, b, 1)$ illustrated in chapter three as well as Poisson Lindley and Zero-truncated Poisson Lindley distributions illustrated in chapter four. This research introduces estimation of aggregate losses based on the results obtained from our multiple state models. This chapter considers three state, four state, five state and six state models of different cancer transitions. Three state model illustrates cancer patients who transit from Healthy state-Leukemia state-Dead states, four state model illustrates patients who transit from Healthy state-Liver cancer state-Colon state-Dead states, five state model illustrates Healthy state-Stomach cancer state-Pharynx state-Colon state-Dead states and six state model illustrates Healthy state-Oesophagus cancer state-Stomach state-Lung state-Kidney state-Dead states. These models are non recovery models as it assumed that once a patient is infected with cancer their transition probability changes tremendous due to the first infection.

5.2 Multiple state models setup

Consider m possible transition states in a multiple state model. Let π be a finite and countable set such that: $\pi = \{1, 2, 3, ..., m\}$ and leNing the set of direct transitions to be W expressed as:

$$W \subseteq (i, j) | i \neq j \quad \text{but} \quad i, j \in \pi$$

The set π, W is called a multiple state model. Let $Z(r)$ be the state of occupancy by individual under study at time r where $r \geq 0$. $\{Z(r); r = 0, 1, 2, 3, ...\}$ is a time-discrete Markov process if for;

$$0 \leq r_0 \leq r_1 \leq r_2 \leq ... \leq r_m$$

and corresponding states

$$i_0, i_1, i_2, ..., i_m \in s$$

with probability

$$pr[Z(r_n) = i_n, Z(r_{n-1}) = i_{n-1}, ..., Z(r_0) = i_0] > 0$$

satisfies the Markov property:

$$pr[Z(r_n) = i_n | Z(r_{n-1}) = i_{n-1}, \ldots, Z(r_0) = i_0] \qquad pr[Z(r_n) = i_n | Z(r_{n-1}) = i_{n-1}]$$

5.2.1 Transition probabilities

Let probability of moving from a given state to the next state in a Markov process by $p_{i,j}(s,r)$ and be define such that: $p_{i,j}(s,r)$ = represents conditional probability of an individual being in state j at time r provided that they were in state i at time s.

$$p_{i,j}(s,r) = pr[Z(r) = j | Z(s) = i] \qquad = \frac{pr[Z(r) = j, Z(s) = i]}{pr[Z(s) = i]}$$

where; $r \geq, s \geq 0$ and $i, j \in M$
The above equation holds if $pr[Z(s) = i] > 0$ else $p_{ij}(s,r) = 0$ hence:

$$p_{i,j}(s,s) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if i=j} \end{cases}$$

This implies that an individual under consideration, can only be in one state at a given time but cannot be in two states at the same time. The transition probabilities have the following properties:
$0 \leq p_{ij}(s,r) \leq 1$
where ; $i, j \in M$

$$\sum_{j \in M} p_{ij}(s,r) = 1$$

This model is build on the assumption that:

* The transition process depends on the length of time interval $[s;r]$

* It also depends on the time s and r when it starts and ends.

Hence the process is classified as time-inhomogeneous.

5.2.2 Chapman-Kolmogorov equation

Definition 5.2.1 (Chapman-Kolmogorov equation). *It states that, a process which begins in state i at time s and is in state j at time r occurs through some arbitrary state $k \in M$ at an arbitrary intermediate* *time R i.e*

$$p_{ij}(s,r) = \sum_{k=1}^{n} pr[Z(s,r) = j, Z(s,r) = k|Z(s) = i] = \sum_{k=1}^{n} \frac{pr[Z(s,r) = j, Z(s,r) = k, Z(s) = i]}{prZ(s) = i}$$

$$= \sum_{k=1}^{n} \frac{pr[Z(s,r) = j|Z(s,r) = k, Z(s) = i] * pr[Z(s,r) = k, Z(s) = i]}{pr[Z(s) = i]}$$

$$= \sum_{k=1}^{n} pr[Z(s,r) = j|Z(s,r) = k, Z(s) = i] * pr[Z(s,r) = k|Z(s) = i]$$

Using Markov property:

$$p_{ij}(s,r) = \sum_{k=1}^{n} pr[Z(s,r) = j|Z(s,r) = k] * pr[Z(s,r) = k|Z(s) = i] = \sum_{k=1}^{n} p_{kj}(R,r)p_{ik}(s,R)$$

$$p_{ij}(s,r) = \sum_{k=1}^{n} p_{ik}(s,R)p_{kj}(R,r) \qquad \textbf{(5.1)}$$

Kolmogorov Forward equation

Equation (5.1) can be transformed to its differential form. Let $r = W_t + \kappa_d$ implying that $R = W_t$ hence equation (5.1) can be rewriNen as:

$$p_{ij}(s,W_t + \kappa_d) = \sum_{k=1}^{n} p_{ik}(s,W_t)p_{kj}(W_t,W_t + \kappa_d)$$

$$p_{ij}(s,W_t + \kappa_d) - p_{ij}(s,W_t) = \sum_{k=1}^{n} p_{ik}(s,W_t)p_{kj}(W_t,W_t + \kappa_d) + p_{ij}(s,W_t)p_{jj}(W_t,W_t + \kappa_d) - p_{ij}(s,W_t)$$

$$= \sum_{k=1}^{n} p_{ik}(s,W_t)p_{kj}(W_t,W_t + \kappa_d) - p_{ij}(s,W_t)[1 - p_{jj}(W_t,W_t + \kappa_d)]$$

$$\lim_{\kappa_d \to 0} \frac{p_{ij}(s,W_t + \kappa_d) - p_{ij}(s,W_t)}{\kappa_d} = \lim_{\kappa_d \to 0} \frac{\sum_{k=1}^{n} p_{ik}(s,W_t)p_{kj}(W_t,W_t + \kappa_d) - p_{ij}(s,W_t)[1 - p_{jj}(W_t,W_t + \kappa_d)]}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{ij}(s,W_t) = \sum_{k=1}^{n} p_{ik}(s,W_t) \lim_{\kappa_d \to 0} \frac{p_{kj}(W_t,W_t + \kappa_d)}{\kappa_d} - p_{ij}(s,W_t) \lim_{\kappa_d \to 0} \frac{[1 - p_{jj}(W_t,W_t + \kappa_d)]}{\kappa_d}$$

$$= \sum_{k=1}^{n} p_{ik}(s,W_t)\Im_{kj} - p_{ij}(s,W_t)\Im_j \qquad \textbf{(5.2)}$$

where : $\lim_{\kappa_d \to 0} \frac{p_{kj}(W_t,W_t+\kappa_d)}{\kappa_d} = \Im_{kj}$ $\lim_{\kappa_d \to 0} \frac{[1-p_{jj}(W_t,W_t+\kappa_d)]}{\kappa_d} = \Im_j$ The Kolmogorov Forward equation is therefore:

$$\frac{\partial}{\partial W_t} p_{ij}(s,W_t) = \sum_{k=1}^{n} p_{ik}(s,W_t)\Im_{kj} - p_{ij}(s,W_t)\Im_i$$

where $\Im_{ij}(W_t)$ represents transition intensity between two states i and j. In a summary we can say that, $\Im_{ij}(W_t)$ represents the change rate of the transition probability p_{ij} in a small time interval, h

$$\Im_{ij}(W_t) = \lim_{\kappa_d \to 0} \frac{p_{ij}(W_t, W_t + \kappa_d)}{\kappa_d} \quad \text{for} \quad i \quad j$$

In our study, we consider and define the transition intensities and probabilities between different cancer states such that:
$p_{ij}(V_A, W_t)$ = The probability that a life aged $V_A + W_t$ and in state j was in state i at age V_A.
$\Im_{V_A+W_t}$ = The transition intensity/rate from state i to state j at age $V_a + W_t$. In the above cases, $i, j = 1, 2, 3, ..., m$ $V_A = 0, 1, 2, 3, ...$ and $0 \leq W_t \leq 1$. It is important to note that we consider constant force of mortality hence we have;

$$\Im_{ij}(V_A + W_t) = \Im_{ij}(V_A) \text{ for } x = 0, 1, 2, 3 \dots \text{ and } 0 \leq W_t \leq 1$$

Hence ;

$$p_{ij}(s, W_t + \kappa_d) = \sum_{k=1}^{n} p_{ik}(s, W_t) p_{kj}(W_t, W_t + \kappa_d) \qquad p_{ij}(V_A, W_t + \kappa_d) = \sum_{k=1}^{n} p_{ik}(V_A, W_t) p_{kj}(W_t, W_t + \kappa_d)$$

where:

$$p_{i,j}(V_A, V_A) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i=j \end{cases}$$

implies that a person can only be in a particular state at a particular time but cannot be in two stages at the same time.

5.3 Three state cancer model

5.3.1 Introduction

This section considers a three state cancer Markov model which in cooperates Healthy state, Leukemia and Dead state. A case where the patients can not move back to any state is considered. This leads to the assumption that a life is not permiNed to enter a state more than one time. The patients who recover are assumed to have been censored from the study. Leukemia was considered for three state model as most of the patient did not transit to any other type of cancer.

5.3.2 Three state Leukemia cancer model

Figure (5.1) represents the three-state model in which we systematically derive the respective Kolmogorov Forward Differential Equation.

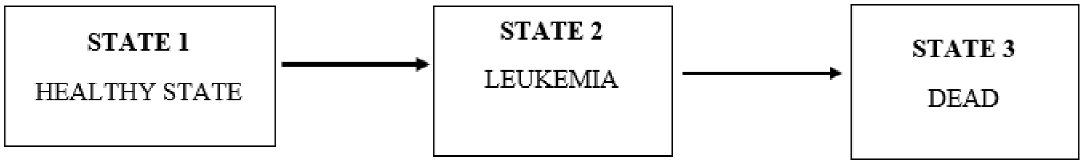


Figure 5.1. Leukemia cancer model

From the figure (5.1) above the transition probability matrix is expressed as:

$$\begin{bmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & 0 \\ 0 & c_1 & d_1 \\ 0 & 0 & 1 \end{bmatrix}$$

Hence the values of interest are; a_1, b_1, c_1 and d_1. Where a_1 represents the transition probability of remaining in healthy state, b_1 represents the transition probability of transiting from healthy state to Leukemia state, c_1 represents the transition probability of remaining in Leukemia state and d_1 represents the transition probability of transiting from Leukemia state to Dead state.

Theorem 5.3.1 (Three state Leukemia model). *The transition probability matrix for this model is:*

$$\begin{bmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) \end{bmatrix} = \begin{bmatrix} e^{-\Im_{12}(V_A)t} & 1-e^{-\Im_{12}(V_A)t} & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1-e^{-\Im_{23}(V_A)t} \\ 0 & 0 & 1 \end{bmatrix}$$

Proof of theorem 5.3.1

Figure 5.1 represents the Three state Leukemia model. The transition intensities and probabilities are derived using Kolmogorov forward equations. Kolmogorov Forward Differential equation is expressed as:

$$p_{ij}(V_A,W_t+\kappa_d) = \sum_{k=1}^{n} p_{ik}(V_A,W_t)p_{kj}(W_t,W_t+\kappa_d)$$

In this model we consider the following: $(i,k,j) = 1,2,3$

When $i = 1$ and $j = 1$

$$\begin{aligned} p_{11}(V_A,W_t+\kappa_d) &= \sum_{k=1}^{3} p_{1k}(V_A,W_t)p_{k1}(W_t,W_t+\kappa_d) \\ &= p_{11}(V_A,W_t)[1-p_{12}(W_t,W_t+\kappa_d)] + p_{12}(V_A,W_t)p_{21}(W_t,W_t+\kappa_d) \\ &\quad + p_{13}(V_A,W_t)p_{31}(W_t,W_t+\kappa_d) \\ &= p_{11}(V_A,W_t)[1-p_{12}(W_t,W_t+\kappa_d)] + p_{12}(V_A,W_t)p_{21}*0 + p_{13}(V_A,W_t)*0 \\ &= p_{11}(V_A,W_t)[1-p_{12}(W_t,W_t+\kappa_d)] \end{aligned} \quad \textbf{(5.3)}$$

Subtracting $p_{11}(V_A, W_t)$ from equation (5.3)

$$p_{11}(V_A, W_t + \kappa_d) - p_{11}(V_A, W_t) = p_{11}(V_A, W_t)[1 - p_{12}(W_t, W_t + \kappa_d)] - p_{11}(V_A, W_t)$$

$$= p_{11}(V_A, W_t)[-p_{12}(W_t, W_t + \kappa_d)]$$

$$\lim_{\kappa_d \to 0} \frac{p_{11}(V_A, W_t + \kappa_d) - p_{11}(V_A, W_t)}{\kappa_d} = p_{11}(V_A, W_t) \lim_{\kappa_d \to 0} \frac{-p_{12}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{11}(V_A, W_t) = p_{11}(V_A, W_t) * -\Im_{12}(V_A) \quad \textbf{(5.4)}$$

When $i = 1$ and $j = 2$

$$p_{12}(V_A, W_t + \kappa_d) = \sum_{k=1}^{3} p_{1k}(V_A, W_t) p_{k2}(W_t, W_t + \kappa_d)$$

$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] + p_{13}(V_A, W_t) p_{32}(W_t, W_t + \kappa_d)$$

$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] \quad \textbf{(5.5)}$$

Subtracting $p_{12}(V_A, W_t)$ from equation (5.5)

$$p_{12}(V_A, W_t + \kappa_d) - p_{12}(V_A, W_t) = p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] - p_{12}(V_A, W_t)$$

$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t) * -p_{23}(W_t, W_t + \kappa_d)$$

$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d)$$

$$\lim_{\kappa_d \to 0} \frac{p_{12}(V_A, W_t + \kappa_d) - p_{12}(V_A, W_t)}{\kappa_d} = p_{11}(V_A, W_t) \lim_{\kappa_d \to 0} \frac{p_{12}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = p_{11}(V_A, W_t) * \Im_{12}(V_A) \quad \textbf{(5.6)}$$

When $i = 2$ and $j = 2$

$$p_{22}(V_A, W_t + \kappa_d) = \sum_{k=1}^{3} p_{2k}(V_A, W_t) p_{k2}(W_t, W_t + \kappa_d)$$

$$= p_{21}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] + p_{23}(V_A, W_t) p_{32}(W_t, W_t + \kappa_d)$$

$$= p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] \quad \textbf{(5.7)}$$

Subtracting $p_{22}(V_A, W_t)$ from equation (5.7)

$$p_{22}(V_A, W_t + \kappa_d) - p_{22}(V_A, W_t) = p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] - p_{22}(V_A, W_t)$$

$$= p_{22}(V_A, W_t) * -p_{23}(W_t, W_t + \kappa_d)$$

$$\lim_{\kappa_d \to 0} \frac{p_{22}(V_A, W_t + \kappa_d) - p_{22}(V_A, W_t)}{\kappa_d} = p_{22}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{23}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = p_{22}(V_A, W_t) * -\Im_{23}(V_A) \quad \textbf{(5.8)}$$

When $i = 2$ and $j = 3$

$$p_{23}(V_A, W_t + \kappa_d) = \sum_{k=1}^{3} p_{2k}(V_A, W_t) p_{k3}(W_t, W_t + \kappa_d)$$
$$= p_{21}(V_A, W_t) p_{13}(W_t, W_t + \kappa_d) + p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d)$$
$$= p_{22}(V_A, W_t) * p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t) \quad \textbf{(5.9)}$$

Subtracting $p_{23}(V_A, W_t)$ from equation (5.9)

$$p_{23}(V_A, W_t + \kappa_d) - p_{23}(V_A, W_t) = p_{22}(V_A, W_t) * p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t) - p_{23}(V_A, W_t)$$
$$= p_{22}(V_A, W_t) * p_{23}(W_t, W_t + \kappa_d)$$
$$\lim_{\kappa_d \to 0} \frac{p_{23}(V_A, W_t + \kappa_d) - p_{23}(V_A, W_t)}{\kappa_d} = p_{22}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{23}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{23}(V_A, W_t) = p_{22}(V_A, W_t) * \Im_{23}(V_A) \quad \textbf{(5.10)}$$

This model is build on the assumption that our observations are done within the interval time $(0, t)$. We solved individual derivatives to aNain the required transition probabilities. Solving for the derivative of equation (5.4) we obtained the transition probability of remaining in healthy state. The derivative of equation (5.4) is expressed as:

$$\frac{\partial}{\partial W_t} p_{11}(V_A, W_t) = p_{11}(V_A, W_t) * -\Im_{12}(V_A) \qquad \frac{\partial}{\partial W_t} \ln p_{11}(V_A, W_t) = -\Im_{12}(V_A)$$
$$\int_0^t \frac{\partial}{\partial W_t} \ln p_{11}(V_A, W_t) \partial W_t = \int_0^t -\Im_{12}(V_A) ds \qquad \ln p_{11}(V_A, W_t) = [-\Im_{12}(V_A) s]_0^t$$
$$\ln p_{11}(V_A, W_t) = [-\Im_{12}(V_A) t] \qquad p_{11}(V_A, W_t) = e^{-\Im_{12}(V_A) t} \quad \textbf{(5.11)}$$

Solving for the derivative of equation (5.6) we obtained the transition probability of moving from healthy state to Leukemia state. The derivative of equation (5.6) is expressed as:

$$\frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = p_{11}(V_A, W_t) * \Im_{12}(V_A) \qquad \frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = e^{-\Im_{12}(V_A) t} * \Im_{12}(V_A)$$
$$\int_0^t \frac{\partial}{\partial W_t} p_{12}(V_A, W_t) \partial W_t = \int_0^t e^{-\Im_{12}(V_A) s} * \Im_{12}(V_A) ds \qquad p_{12}(V_A, W_t) = \Im_{12}(V_A) \int_0^t e^{-\Im_{12}(V_A) s} ds$$
$$p_{12}(V_A, W_t) = \left[\frac{\Im_{12}(V_A) * e^{-\Im_{12}(V_A) s}}{-\Im_{12}(V_A)} \right]_0^t \qquad p_{12}(V_A, W_t) = 1 - e^{-\Im_{12}(V_A) t} \quad \textbf{(5.12)}$$

Solving for the derivative of equation (5.8) we obtained the transition probability of remaining in Leukemia state. The derivative of equation (5.8) is expressed as:

$$\frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = p_{22}(V_A, W_t) * -\Im_{23}(V_A) \qquad \frac{1}{p_{22}(V_A, W_t)} \frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = -\Im_{23}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} \ln p_{22}(V_A, W_t)\partial W_t = \int_0^t -\Im_{23}(V_A)ds \qquad \int_0^t \partial \ln p_{22}(V_A, W_t) = \int_0^t -\Im_{23}(V_A)ds$$

$$\ln p_{22}(V_A, W_t) = [-\Im_{23}(V_A)s]^t_0 \qquad p_{22}(V_A, W_t) = e^{-\Im_{23}(V_A)t} \tag{5.13}$$

Solving for the derivative of equation (5.10) we obtained the transition probability of moving from Leukemia state to Dead state. The derivative of equation (5.10) is expressed as:

$$\frac{\partial}{\partial W_t} p_{23}(V_A, W_t) = e^{-\Im_{23}(V_A)t} * \Im_{23}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t} p_{23}(V_A, W_t)\partial W_t = \int_0^t e^{-\Im_{23}(V_A)s} * \Im_{23}(V_A)ds$$

$$p_{23}(V_A, W_t) = \Im_{23}(V_A) \int_0^t e^{-\Im_{23}(V_A)s} ds \qquad p_{23}(V_A, W_t) = \left[\frac{\Im_{23}(V_A) * e^{-\Im_{23}(V_A)s}}{-\Im_{23}(V_A)}\right]_0^t$$

$$p_{23}(V_A, W_t) = -e^{-\Im_{23}(V_A)t} + 1 \qquad p_{23}(V_A, W_t) = 1 - e^{-\Im_{23}(V_A)t} \tag{5.14}$$

In this model state 3 is an absorbing state hence the transition probabilities are defined as: $p_{31}(V_A, W_t) = 0$, $p_{32}(V_A, W_t) = 0$, $p_{31}(V_A, W_t) = 1$ The transition probability matrix for this model is :

$$\begin{bmatrix} p_{11}(V_A, W_t) & p_{12}(V_A, W_t) & p_{13}(V_A, W_t) \\ p_{21}(V_A, W_t) & p_{22}(V_A, W_t) & p_{23}(V_A, W_t) \\ p_{31}(V_A, W_t) & p_{32}(V_A, W_t) & p_{33}(V_A, W_t) \end{bmatrix} = \begin{bmatrix} e^{-\Im_{12}(V_A)t} & 1 - e^{-\Im_{12}(V_A)t} & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1 - e^{-\Im_{23}(V_A)t} \\ 0 & 0 & 1 \end{bmatrix}$$

5.4 Four state cancer model

5.4.1 Introduction

In this section four state cancer Markov model which in cooperates Healthy state, Liver cancer state,Colon state and Dead state is derived. A case where patients can not move back to any state is considered. This leads to the assumption that a life cannot enter a state more than once. The patients who recover are assumed to have been censored from the study.

5.4.2 Four state Liver cancer-Colon model

Figure (5.2) represents the four-state model in which we systematically derive the respective Kolmogorov Forward Differential Equation.

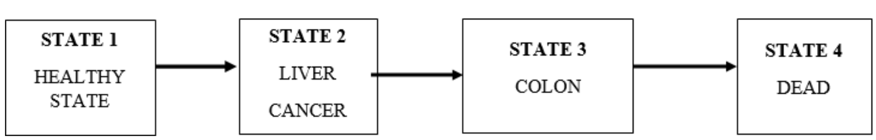


Figure 5.2. Liver cancer-Colon model

From the figure (5.2) above the transition probability matrix is expressed as:

$$\begin{pmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) \end{pmatrix} = \begin{pmatrix} a_2 & b_2 & 0 & 0 \\ 0 & c_2 & d_2 & 0 \\ 0 & 0 & e_2 & f_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Hence the values of interest are; a_2, b_2, c_2, d_2, e_2 and f_2. Where a_2 represents transition probability of remaining in healthy state, b_2 represents the transition probability of moving from Healthy state to Liver cancer state, c_2 represents the transition probability of remaining in Liver cancer state, d_2 represents the transition probability of moving from Liver cancer state to Colon cancer state and e_2 represents the transition probability of remaining in Colon cancer state and f_2 is the transition probability of moving from Colon cancer state to Dead state.

Theorem 5.4.1 (Four state Liver-Colon cancer model). *The transition probability matrix for this model is:*

$$\begin{pmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) \end{pmatrix} = \begin{pmatrix} e^{-\Im_{12}(V_A)t} & 1-e^{-\Im_{12}(V_A)t} & 0 & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1-e^{-\Im_{23}(V_A)t} & 0 \\ 0 & 0 & e^{-\Im_{34}(V_A)t} & 1-e^{-\Im_{34}(V_A)t} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Proof of theorem 5.4.1

Figure (5.2) represents the Four state Liver-Colon cancer model. The transition intensities and probabilities are derived using Kolmogorov forward equations. Kolmogorov Forward Differential equation is expressed as:

$$p_{ij}(V_A, W_t + \kappa_d) = \sum_{k=1}^{n} p_{ik}(V_A, W_t) p_{kj}(W_t, W_t + \kappa_d)$$

In this model we consider the following: $(i, j, k) = 1, 2, 3, 4$

When $i = 1$ and $j = 1$

$$p_{11}(V_A, W_t + \kappa_d) = \sum_{k=1}^{4} p_{1k}(V_A, W_t) p_{k1}(W_t, W_t + \kappa_d)$$
$$+ p_{13}(V_A, W_t) p_{31}(W_t, W_t + \kappa_d) + p_{14}(V_A, W_t) p_{41}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t)[1 - p_{12}(W_t, W_t + \kappa_d)] + p_{12}(V_A, W_t) p_{21}(W_t, W_t + \kappa_d)$$
$$+ p_{13}(V_A, W_t) p_{31}(W_t, W_t + \kappa_d) + p_{14}(V_A, W_t) p_{41}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t)[1 - p_{12}(W_t, W_t + \kappa_d)] \quad \textbf{(5.15)}$$

Subtracting $p_{11}(V_A, W_t)$ from equation (5.15)

$$p_{11}(V_A, W_t + \kappa_d) - p_{11}(V_A, W_t) = p_{11}(V_A, W_t)[1 - p_{12}(W_t, W_t + \kappa_d)] - p_{11}(V_A, W_t)$$
$$= p_{11}(V_A, W_t)[-p_{12}(W_t, W_t + \kappa_d)]$$
$$\lim_{\kappa_d \to 0} \frac{p_{11}(V_A, W_t + \kappa_d) - p_{11}(V_A, W_t)}{\kappa_d} = p_{11}(V_A, W_t) \lim_{\kappa_d \to 0} \frac{-p_{12}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{11}(V_A, W_t) = p_{11}(V_A, W_t) * -\Im_{12}(V_A) \quad \textbf{(5.16)}$$

When $i = 1$ and $j = 2$

$$p_{12}(V_A, W_t + \kappa_d) = \sum_{k=1}^{4} p_{1k}(V_A, W_t) p_{k2}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)]$$
$$+ p_{13}(V_A, W_t) p_{32}(W_t, W_t + \kappa_d) + p_{14}(V_A, W_t) p_{42}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] \quad \textbf{(5.17)}$$

Subtracting $p_{12}(V_A, W_t)$ from equation (5.17)

$$p_{12}(V_A, W_t + \kappa_d) - p_{12}(V_A, W_t) = p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)]$$
$$- p_{12}(V_A, W_t)$$
$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t) * -p_{23}(W_t, W_t + \kappa_d)$$
$$\lim_{\kappa_d \to 0} \frac{p_{12}(V_A, W_t + \kappa_d) - p_{12}(V_A, W_t)}{\kappa_d} = p_{11}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{12}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = p_{11}(V_A, W_t) * \Im_{12}(V_A) \quad \textbf{(5.18)}$$

When $i = 2$ and $j = 2$

$$p_{22}(V_A, W_t + \kappa_d) = \sum_{k=1}^{4} p_{2k}(V_A, W_t) p_{k2}(W_t, W_t + \kappa_d)$$
$$= p_{21}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)]$$
$$+ p_{23}(V_A, W_t) p_{32}(W_t, W_t + \kappa_d) + p_{24}(V_A, W_t) p_{42}(W_t, W_t + \kappa_d)$$
$$= p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] \quad \textbf{(5.19)}$$

Subtracting $p_{22}(V_A, W_t)$ from equation (5.19)

$$p_{22}(V_A, W_t + \kappa_d) - p_{22}(V_A, W_t) = p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] - p_{22}(V_A, W_t)$$
$$= p_{22}(V_A, W_t) * -p_{23}(W_t, W_t + \kappa_d)$$
$$\lim_{\kappa_d \to 0} \frac{p_{22}(V_A, W_t + \kappa_d) - p_{22}(V_A, W_t)}{\kappa_d} = p_{22}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{23}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = p_{22}(V_A, W_t) * -\Im_{23}(V_A) \quad \textbf{(5.20)}$$

When $i = 2$ and $j = 3$

$$p_{23}(V_A, W_t + \kappa_d) = \sum_{k=1}^{4} p_{2k}(V_A, W_t) p_{k3}(W_t, W_t + \kappa_d)$$
$$= p_{21}(V_A, W_t) p_{13}(W_t, W_t + \kappa_d) + p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d)$$
$$+ p_{23}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d) + p_{24}(V_A, W_t) p_{43}(W_t, W_t + \kappa_d)$$
$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d)$$
$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] \quad \textbf{(5.21)}$$

Subtracting $p_{23}(V_A, W_t)$ from equation (5.21)

$$p_{23}(V_A, W_t + \kappa_d) - p_{23}(V_A, W_t) = p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)]$$
$$- p_{23}(V_A, W_t)$$
$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)[-p_{34}(W_t, W_t + \kappa_d)]$$
$$\lim_{\kappa_d \to 0} \frac{p_{23}(V_A, W_t + \kappa_d) - p_{23}(V_A, W_t)}{\kappa_d} = p_{22}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{23}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{23}(V_A, W_t) = p_{22}(V_A, W_t) * \Im_{23}(V_A) \quad \textbf{(5.22)}$$

When $i = 3$ and $j = 3$

$$p_{33}(V_A, W_t + \kappa_d) = \sum_{k=1}^{4} p_{3k}(V_A, W_t) p_{k3}(W_t, W_t + \kappa_d)$$
$$= p_{31}(V_A, W_t) p_{13}(W_t, W_t + \kappa_d) + p_{32}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d)$$
$$+ p_{33}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t) p_{43}(W_t, W_t + \kappa_d)$$
$$= p_{33}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] \quad \textbf{(5.23)}$$

Subtracting $p_{33}(V_A, W_t)$ from equation (5.23)

$$p_{33}(V_A, W_t + \kappa_d) - p_{33}(V_A, W_t) = p_{33}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] - p_{33}(V_A, W_t)$$
$$= p_{33}(V_A, W_t)[-p_{34}(W_t, W_t + \kappa_d)]$$
$$\lim_{\kappa_d \to 0} \frac{p_{33}(V_A, W_t + \kappa_d) - p_{33}(V_A, W_t)}{\kappa_d} = p_{33}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{34}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{33}(V_A, W_t) = p_{33}(V_A, W_t) * -\Im_{34}(V_A) \quad \textbf{(5.24)}$$

When $i = 3$ and $j = 4$

$$p_{34}(V_A, W_t + \kappa_d) = \sum_{k=1}^{4} p_{3k}(V_A, W_t) p_{k4}(W_t, W_t + \kappa_d)$$

$$= p_{31}(V_A, W_t) p_{14}(W_t, W_t + \kappa_d) + p_{32}(V_A, W_t) p_{24}(W_t, W_t + \kappa_d) + p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t) p_{44}(W_t, W_t + \kappa_d)$$

$$= p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t) \tag{5.25}$$

Subtracting $p_{34}(V_A, W_t)$ from equation (5.25)

$$p_{34}(V_A, W_t + \kappa_d) - p_{34}(V_A, W_t) = p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t) - p_{34}(V_A, W_t)$$

$$= p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d)$$

$$\lim_{\kappa_d \to 0} \frac{p_{34}(V_A, W_t + \kappa_d) - p_{34}(V_A, W_t)}{\kappa_d} = p_{33}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{34}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{34}(V_A, W_t) = p_{33}(V_A, W_t) * \Im_{34}(V_A) \tag{5.26}$$

This model is build on the assumption that our observations are done within the interval time $(0, t)$. We solved individual derivatives to aNain the required transition probabilities. Solving for the derivative of equation (5.16) we obtained the transition probability of remaining in healthy state. The derivative of equation (5.16) is expressed as:

$$\frac{\partial}{\partial W_t} p_{11}(V_A, W_t) = p_{11}(V_A, W_t) * -\Im_{12}(V_A) \qquad \frac{1}{p_{11}(V_A, W_t)} \frac{\partial}{\partial W_t} p_{11}(V_A, W_t) = -\Im_{12}(V_A)$$

$$\frac{\partial}{\partial W_t} \ln p_{11}(V_A, W_t) = -\Im_{12}(V_A) \qquad \int_0^t \partial \ln p_{11}(V_A, W_t) = \int_0^t -\Im_{12}(V_A) ds$$

$$\ln p_{11}(V_A, W_t) = [-\Im_{12}(V_A) t] \qquad p_{11}(V_A, W_t) = e^{-\Im_{12}(V_A) t} \tag{5.27}$$

Solving for the derivative of equation (5.18) we obtained the transition probability of transiting from healthy state to Liver cancer state. The derivative of equation (5.18) is expressed as:

$$\frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = p_{11}(V_A, W_t) * \Im_{12}(V_A) \qquad \frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = e^{-\Im_{12}(V_A) t} * \Im_{12}(V_A)$$

$$\int_0^t \partial p_{12}(V_A, W_t) = \int_0^t e^{-\Im_{12}(V_A) s} * \Im_{12}(V_A) ds \qquad p_{12}(V_A, W_t) = \Im_{12}(V_A) \int_0^t e^{-\Im_{12}(V_A) s} ds$$

$$p_{12}(V_A, W_t) = -e^{-\Im_{12}(V_A) t} + 1 \qquad p_{12}(V_A, W_t) = 1 - e^{-\Im_{12}(V_A) t} \tag{5.28}$$

Solving for the derivative of equation (5.20) we obtained the transition probability of remaining in Liver cancer state. The derivative of equation (5.20) is expressed as:

$$\frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = p_{22}(V_A, W_t) * -\Im_{23}(V_A) \qquad \frac{1}{p_{22}(V_A, W_t)} \frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = -\Im_{23}(V_A)$$

$$\frac{\partial}{\partial W_t} \ln p_{22}(V_A, W_t) = -\Im_{23}(V_A) \qquad \int_0^t \partial \ln p_{22}(V_A, W_t) = \int_0^t -\Im_{23}(V_A) ds$$

$$\ln p_{22}(V_A, W_t) = [-\Im_{23}(V_A) t] \qquad p_{22}(V_A, W_t) = e^{-\Im_{23}(V_A) t} \tag{5.29}$$

Solving for the derivative of equation (5.22) we obtained the transition probability of moving from Liver cancer state to Colon cancer state. The derivative of equation (5.22) is expressed as:

$$\frac{\partial}{\partial W_t} p_{23}(V_A,W_t) = p_{22}(V_A,W_t) * \Im_{23}(V_A) \qquad \frac{\partial}{\partial W_t} p_{23}(V_A,W_t) = e^{-\Im_{23}(V_A)t} * \Im_{23}(V_A)$$

$$\int_0^t \partial p_{23}(V_A,W_t) = \int_0^t e^{-\Im_{23}(V_A)s} * \Im_{23}(V_A)ds \qquad p_{23}(V_A,W_t) = \Im_{23}(V_A)\int_0^t e^{-\Im_{23}(V_A)s}ds$$

$$p_{23}(V_A,W_t) = \left[\frac{\Im_{23}(V_A) * e^{-\Im_{23}(V_A)s}}{-\Im_{23}(V_A)}\right]_0^t \qquad p_{23}(V_A,W_t) = 1 - e^{-\Im_{23}(V_A)t} \tag{5.30}$$

Solving for the derivative of equation (5.24) we obtained the transition probability of remaining in colon cancer state. The derivative of equation (5.24) is expressed as:

$$\frac{\partial}{\partial W_t} p_{33}(V_A,W_t) = p_{33}(V_A,W_t) * -\Im_{34}(V_A) \qquad \frac{1}{p_{33}(V_A,W_t)}\frac{\partial}{\partial W_t} p_{33}(V_A,W_t) = -\Im_{34}(V_A)$$

$$\frac{\partial}{\partial W_t}\ln p_{33}(V_A,W_t) = -\Im_{34}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t}\ln p_{33}(V_A,W_t)\partial W_t = \int_0^t -\Im_{34}(V_A)ds$$

$$\ln p_{33}(V_A,W_t) = [-\Im_{34}(V_A)t] \qquad p_{33}(V_A,W_t) = e^{-\Im_{34}(V_A)t} \tag{5.31}$$

Solving for the derivative of equation (5.26) we obtained the transition probability of moving from Colon cancer state to Dead state. The derivative of equation (5.26) is expressed as:

$$\frac{\partial}{\partial W_t} p_{34}(V_A,W_t) = p_{33}(V_A,W_t) * \Im_{34}(V_A) \qquad \frac{\partial}{\partial W_t} p_{34}(V_A,W_t) = e^{-\Im_{34}(V_A)t} * \Im_{34}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{34}(V_A,W_t)\partial W_t = \int_0^t e^{-\Im_{34}(V_A)s} * \Im_{34}(V_A)ds \qquad p_{34}(V_A,W_t) = \Im_{34}(V_A)\int_0^t e^{-\Im_{34}(V_A)s}ds$$

$$p_{34}(V_A,W_t) = \left[\frac{\Im_{34}(V_A) * e^{-\Im_{34}(V_A)s}}{-\Im_{34}(V_A)}\right]_0^t \qquad p_{34}(V_A,W_t) = 1 - e^{-\Im_{34}(V_A)t} \tag{5.32}$$

In this model state 4 is an absorbing state hence the transition probabilities are defined as: $p_{41}(V_A,W_t) = 0$, $p_{42}(V_A,W_t) = 0$, $p_{43}(V_A,W_t) = 0$, $p_{44}(V_A,W_t) = 1$. The transition probability matrix for this model is :

$$\begin{bmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) \end{bmatrix} = \begin{bmatrix} e^{-\Im_{12}(V_A)t} & 1-e^{-\Im_{12}(V_A)t} & 0 & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1-e^{-\Im_{23}(V_A)t} & 0 \\ 0 & 0 & e^{-\Im_{34}(V_A)t} & 1-e^{-\Im_{34}(V_A)t} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

5.5 Five state cancer model

5.5.1 Introduction

In this section five state cancer Markov model which in cooperates Healthy state, Stomach cancer state, Pharynx state, colon state and Dead state is derived. A case where the patients can not transit back to any state is considered. This leads to the assumption that a patient cannot enter a state more than once. The patients who recover are assumed to have been censored from the study.

5.5.2 Five state Stomach cancer-Pharynx-Colon model

Figure (5.3) represents the five-state model in which we systematically derive the respective Kolmogorov Forward Differential Equation.

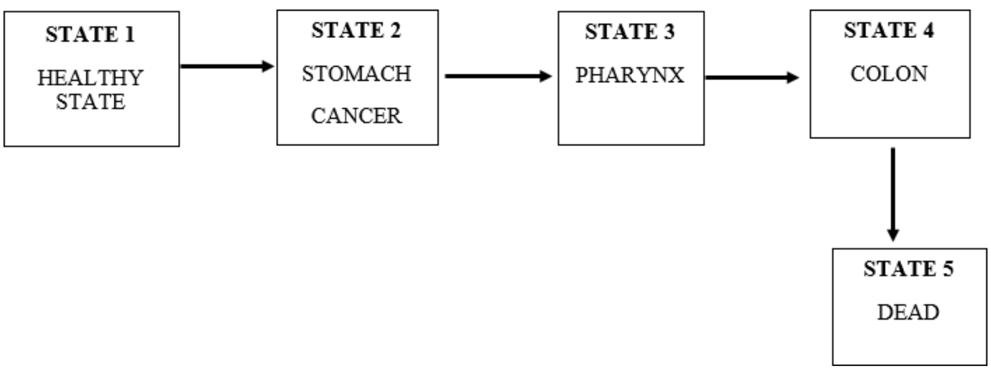


Figure 5.3. Stomach cancer-Pharynx-Colon model

From the figure (5.3)above the transition probability matrix is expressed as:

$$\begin{pmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) & p_{15}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) & p_{25}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) & p_{35}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) & p_{45}(V_A,W_t) \\ p_{51}(V_A,W_t) & p_{52}(V_A,W_t) & p_{53}(V_A,W_t) & p_{54}(V_A,W_t) & p_{55}(V_A,W_t) \end{pmatrix} = \begin{pmatrix} a_3 & b_3 & 0 & 0 & 0 \\ 0 & c_3 & d_3 & 0 & 0 \\ 0 & 0 & e_3 & f_3 & 0 \\ 0 & 0 & 0 & g_3 & h_3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Hence the transition probabilities of interest are; a_3, b_3, c_3, d_3, e_3, f_3, g_3 and h_3. Where a_3 represents transition probability of remaining in healthy state, b_3 represents the transition probability of moving from Healthy state to Stomach cancer state, c_3 represents the transition probability of remaining in Stomach cancer state, d_3 represents the transition probability of transiting from Stomach cancer state to Pharynx cancer state, e_3 represents the transition probability of remaining in Pharynx cancer state, f_3 is the transition probability of transiting from Pharynx cancer state to Colon cancer state, g_3 represents the transition probability of remaining in Colon cancer state and h_3 is the transition probability of moving from Colon cancer state to Dead cancer state.

Theorem 5.5.1 (Five state Stomach-Pharynx-Colon cancer model). *The transition probability matrix for this model is:*

$$\begin{bmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) & p_{15}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) & p_{25}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) & p_{35}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) & p_{45}(V_A,W_t) \\ p_{51}(V_A,W_t) & p_{52}(V_A,W_t) & p_{53}(V_A,W_t) & p_{54}(V_A,W_t) & p_{55}(V_A,W_t) \end{bmatrix} =$$

$$\begin{bmatrix} e^{-\Im_{12}(V_A)t} & 1-e^{-\Im_{12}(V_A)t} & 0 & 0 & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1-e^{-\Im_{23}(V_A)t} & 0 & 0 \\ 0 & 0 & e^{-\Im_{34}(V_A)t} & 1-e^{-\Im_{34}(V_A)t} & 0 \\ 0 & 0 & 0 & e^{-\Im_{45}(V_A)t} & 1-e^{-\Im_{45}(V_A)t} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Proof of theorem 5.5.1

Figure 5.3 represents the Five state Stomach-Pharynx-Colon cancer model. The transition intensities and probabilities are derived using Kolmogorov forward equations. Kolmogorov Forward Differential equation is expressed as:

$$p_{ij}(V_A,W_t+\kappa_d)=\sum_{k=1}^{n} p_{ik}(V_A,W_t)p_{kj}(W_t,W_t+\kappa_d)$$

In this model we consider the following: $(i,k,j)=1,2,3,4,5$

When $i=1$ and $j=1$

$$\begin{aligned} p_{11}(V_A,W_t+\kappa_d) &= \sum_{k=1}^{5} p_{1k}(V_A,W_t)p_{k1}(W_t,W_t+\kappa_d) \\ &= p_{11}(V_A,W_t)[1-p_{12}(W_t,W_t+\kappa_d)]+p_{12}(V_A,W_t)p_{21}(W_t,W_t+\kappa_d) \\ &\quad + p_{13}(V_A,W_t)p_{31}(W_t,W_t+\kappa_d)+p_{14}(V_A,W_t)p_{41}(W_t,W_t+\kappa_d) \\ &\quad + p_{15}(V_A,W_t)p_{51}(W_t,W_t+\kappa_d) \\ &= p_{11}(V_A,W_t)[1-p_{12}(W_t,W_t+\kappa_d)] \end{aligned} \quad \textbf{(5.33)}$$

Subtracting $p_{11}(V_A,W_t)$ from equation (5.33)

$$\begin{aligned} p_{11}(V_A,W_t+\kappa_d)-p_{11}(V_A,W_t) &= p_{11}(V_A,W_t)[1-p_{12}(W_t,W_t+\kappa_d)]-p_{11}(V_A,W_t) \\ &= p_{11}(V_A,W_t)[1-p_{12}(W_t,W_t+\kappa_d)-1] \\ &= p_{11}(V_A,W_t)[-p_{12}(W_t,W_t+\kappa_d)] \end{aligned}$$

$$\lim_{\kappa_d\to 0}\frac{p_{11}(V_A,W_t+\kappa_d)-p_{11}(V_A,W_t)}{\kappa_d}=p_{11}(V_A,W_t)\lim_{\kappa_d\to 0}\frac{-p_{12}(W_t,W_t+\kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t}p_{11}(V_A,W_t)=p_{11}(V_A,W_t)*-\Im_{12}(V_A) \quad \textbf{(5.34)}$$

When $i = 1$ and $j = 2$

$$p_{12}(V_A, W_t + \kappa_d) = \sum_{k=1}^{5} p_{1k}(V_A, W_t) p_{k2}(W_t, W_t + \kappa_d)$$

$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)]$$
$$+ p_{13}(V_A, W_t) p_{32}(W_t, W_t + \kappa_d) + p_{14}(V_A, W_t) p_{42}(W_t, W_t + \kappa_d)$$
$$+ p_{15}(V_A, W_t) p_{52}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] \quad \textbf{(5.35)}$$

Subtracting $p_{12}(V_A, W_t)$ from equation (5.35)

$$p_{12}(V_A, W_t + \kappa_d) - p_{12}(V_A, W_t) = p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)$$
$$[1 - p_{23}(W_t, W_t + \kappa_d)] - p_{12}(V_A, W_t)]$$
$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)$$
$$* - p_{23}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t) * p_{12}(W_t, W_t + \kappa_d)$$

$$\lim_{\kappa_d \to 0} \frac{p_{12}(V_A, W_t + \kappa_d) - p_{12}(V_A, W_t)}{\kappa_d} = p_{11}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{12}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = p_{11}(V_A, W_t) * \Im_{12}(V_A) \quad \textbf{(5.36)}$$

When $i = 2$ and $j = 2$

$$p_{22}(V_A, W_t + \kappa_d) = \sum_{k=1}^{5} p_{2k}(V_A, W_t) p_{k2}(W_t, W_t + \kappa_d)$$

$$= p_{21}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)]$$
$$+ p_{23}(V_A, W_t) p_{32}(W_t, W_t + \kappa_d) + p_{24}(V_A, W_t) p_{42}(W_t, W_t + \kappa_d)$$
$$p_{25}(V_A, W_t) p_{52}(W_t, W_t + \kappa_d)$$
$$= p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] \quad \textbf{(5.37)}$$

Subtracting $p_{22}(V_A, W_t)$ from equation (5.37)

$$p_{22}(V_A, W_t + \kappa_d) - p_{22}(V_A, W_t) = p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] - p_{22}(V_A, W_t)$$
$$= p_{22}(V_A, W_t) * - p_{23}(W_t, W_t + \kappa_d)$$

$$\lim_{\kappa_d \to 0} \frac{p_{22}(V_A, W_t + \kappa_d) - p_{22}(V_A, W_t)}{\kappa_d} = p_{22}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{23}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = p_{22}(V_A, W_t) * - \Im_{23}(V_A) \quad \textbf{(5.38)}$$

When $i = 2$ and $j = 3$

$$p_{23}(V_A, W_t + \kappa_d) = \sum_{k=1}^{5} p_{2k}(V_A, W_t) p_{k3}(W_t, W_t + \kappa_d)$$

$$= p_{21}(V_A, W_t) p_{13}(W_t, W_t + \kappa_d) + p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d) + p_{24}(V_A, W_t) p_{43}(W_t, W_t + \kappa_d) + p_{25}(V_A, W_t) p_{53}(W_t, W_t + \kappa_d)$$

$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d)$$

$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] \quad \textbf{(5.39)}$$

Subtracting $p_{23}(V_A, W_t)$ from equation (5.39)

$$p_{23}(V_A, W_t + \kappa_d) - p_{23}(V_A, W_t) = p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] - p_{23}(V_A, W_t)$$

$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)[-p_{34}(W_t, W_t + \kappa_d)]$$

$$\lim_{\kappa_d \to 0} \frac{p_{23}(V_A, W_t + \kappa_d) - p_{23}(V_A, W_t)}{\kappa_d} = p_{22}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{23}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{23}(V_A, W_t) = p_{22}(V_A, W_t) * \Im_{23}(V_A) \quad \textbf{(5.40)}$$

When $i = 3$ and $j = 3$

$$p_{33}(V_A, W_t + \kappa_d) = \sum_{k=1}^{5} p_{3k}(V_A, W_t) p_{k3}(W_t, W_t + \kappa_d)$$

$$= p_{31}(V_A, W_t) p_{13}(W_t, W_t + \kappa_d) + p_{32}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{33}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t) p_{43}(W_t, W_t + \kappa_d) + p_{35}(V_A, W_t) p_{53}(W_t, W_t + \kappa_d)$$

$$= p_{33}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] \quad \textbf{(5.41)}$$

Subtracting $p_{33}(V_A, W_t)$ from equation (5.41)

$$p_{33}(V_A, W_t + \kappa_d) - p_{33}(V_A, W_t) = p_{33}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] - p_{33}(V_A, W_t)$$

$$= p_{33}(V_A, W_t)[-p_{34}(W_t, W_t + \kappa_d)]$$

$$\lim_{\kappa_d \to 0} \frac{p_{33}(V_A, W_t + \kappa_d) - p_{33}(V_A, W_t)}{\kappa_d} = p_{33}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{34}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{33}(V_A, W_t) = p_{33}(V_A, W_t) * -\Im_{34}(V_A) \quad \textbf{(5.42)}$$

When $i = 3$ and $j = 4$

$$p_{34}(V_A, W_t + \kappa_d) = \sum_{k=1}^{5} p_{3k}(V_A, W_t) p_{k4}(W_t, W_t + \kappa_d)$$

$$= p_{31}(V_A, W_t) p_{14}(W_t, W_t + \kappa_d) + p_{32}(V_A, W_t) p_{24}(W_t, W_t + \kappa_d)$$
$$+ p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t) p_{44}(W_t, W_t + \kappa_d)$$
$$+ p_{35}(V_A, W_t) p_{54}(W_t, W_t + \kappa_d)$$
$$= p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t)[1 - p_{45}(W_t, W_t + \kappa_d)] \quad \textbf{(5.43)}$$

Subtracting $p_{34}(V_A, W_t)$ from equation (5.43)

$$p_{34}(V_A, W_t + \kappa_d) - p_{34}(V_A, W_t) = p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t)$$
$$[1 - p_{45}(W_t, W_t + \kappa_d)] - p_{34}(V_A, W_t)$$
$$= p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d) + p_{34}(V_A, W_t)$$
$$[-p_{45}(W_t, W_t + \kappa_d)]$$
$$= p_{33}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d)$$

$$\lim_{\kappa_d \to 0} \frac{p_{34}(V_A, W_t + \kappa_d) - p_{34}(V_A, W_t)}{\kappa_d} = p_{33}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{34}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{34}(V_A, W_t) = p_{33}(V_A, W_t) * \Im_{34}(V_A) \quad \textbf{(5.44)}$$

When $i = 4$ and $j = 4$

$$p_{44}(V_A, W_t + \kappa_d) = \sum_{k=1}^{5} p_{4k}(V_A, W_t) p_{k4}(W_t, W_t + \kappa_d)$$

$$= p_{41}(V_A, W_t) p_{14}(W_t, W_t + \kappa_d) + p_{42}(V_A, W_t) p_{24}(W_t, W_t + \kappa_d)$$
$$+ p_{43}(V_A, W_t) p_{34}(W_t, W_t + \kappa_d) + p_{44}(V_A, W_t) p_{44}(W_t, W_t + \kappa_d)$$
$$+ p_{45}(V_A, W_t) p_{54}(W_t, W_t + \kappa_d)$$
$$= p_{44}(V_A, W_t)[1 - p_{45}(W_t, W_t + \kappa_d)] \quad \textbf{(5.45)}$$

Subtracting $p_{44}(V_A, W_t)$ from equation (5.45)

$$p_{44}(V_A, W_t + \kappa_d) = p_{44}(V_A, W_t)[1 - p_{45}(W_t, W_t + \kappa_d)]$$
$$= p_{44}(V_A, W_t)[-p_{45}(W_t, W_t + \kappa_d)]$$

$$\lim_{\kappa_d \to 0} \frac{p_{44}(V_A, W_t + \kappa_d) - p_{44}(V_A, W_t)}{\kappa_d} = p_{44}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{45}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{44}(V_A, W_t) = p_{44}(V_A, W_t) * -\Im_{45}(V_A) \quad \textbf{(5.46)}$$

When $i = 4$ and $j = 5$

$$\begin{aligned} p_{45}(V_A, W_t + \kappa_d) &= \sum_{k=1}^{5} p_{4k}(V_A, W_t) p_{k5}(W_t, W_t + \kappa_d) \\ &= p_{41}(V_A, W_t) p_{15}(W_t, W_t + \kappa_d) + p_{42}(V_A, W_t) p_{25}(W_t, W_t + \kappa_d) \\ &\quad + p_{43}(V_A, W_t) p_{35}(W_t, W_t + \kappa_d) + p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d) \\ &\quad + p_{45}(V_A, W_t) p_{55}(W_t, W_t + \kappa_d) \\ &= p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d) + p_{45}(V_A, W_t) \end{aligned}$$

(5.47)

Subtracting $p_{45}(V_A, W_t)$ from equation (5.47)

$$\begin{aligned} p_{45}(V_A, W_t + \kappa_d) - p_{45}(V_A, W_t) &= p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d) + p_{45}(V_A, W_t) - p_{45}(V_A, W_t) \\ &= p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d) \end{aligned}$$

$$\lim_{\kappa_d \to 0} \frac{p_{45}(V_A, W_t + \kappa_d) - p_{45}(V_A, W_t)}{\kappa_d} = p_{44}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{45}(W_t, W_t + \kappa_d)}{\kappa_d}$$

$$\frac{\partial}{\partial W_t} p_{45}(V_A, W_t) = p_{44}(V_A, W_t) * \Im_{45}(V_A) \qquad \textbf{(5.48)}$$

This model is build on the assumption that our observations are done within the interval time $(0,\ t)$. We solved individual derivatives to aNain the required transition probabilities. Solving for the derivative of equation (5.34) we obtained the transition probability of remaining in healthy state. The derivative of equation (5.34) is expressed as:

$$\frac{\partial}{\partial W_t} p_{11}(V_A, W_t) = p_{11}(V_A, W_t) * -\Im_{12}(V_A) \qquad \frac{1}{p_{11}(V_A, W_t)} \frac{\partial}{\partial W_t} p_{11}(V_A, W_t) = -\Im_{12}(V_A)$$

$$\frac{\partial}{\partial W_t} \ln p_{11}(V_A, W_t) = -\Im_{12}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t} \ln p_{11}(V_A, W_t) \partial W_t = \int_0^t -\Im_{12}(V_A) ds$$

$$\ln p_{11}(V_A, W_t) = [-\Im_{12}(V_A) t] \qquad p_{11}(V_A, W_t) = e^{-\Im_{12}(V_A) t} \qquad \textbf{(5.49)}$$

Solving for the derivative of equation (5.36) we obtained the transition probability of transiting from healthy state to Stomach cancer state. The derivative of equation (5.36) is expressed as:

$$\frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = p_{11}(V_A, W_t) * \Im_{12}(V_A) \qquad \frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = e^{-\Im_{12}(V_A) t} * \Im_{12}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{12}(V_A, W_t) \partial W_t = \int_0^t e^{-\Im_{12}(V_A) s} * \Im_{12}(V_A) ds \qquad p_{12}(V_A, W_t) = \Im_{12}(V_A) \int_0^t e^{-\Im_{12}(V_A) s} ds$$

$$p_{12}(V_A, W_t) = \left[\frac{\Im_{12}(V_A) * e^{-\Im_{12}(V_A) s}}{-\Im_{12}(V_A)} \right]_0^t \qquad p_{12}(V_A, W_t) = 1 - e^{-\Im_{12}(V_A) t} \qquad \textbf{(5.50)}$$

Solving for the derivative of equation (5.38) we obtained the transition probability of remaining in stomach cancer state. The derivative of equation (5.38) is expressed as:

$$\frac{\partial}{\partial W_t} p_{22}(V_A,W_t) = p_{22}(V_A,W_t) * -\Im_{23}(V_A) \qquad \frac{1}{p_{22}(V_A,W_t)}\frac{\partial}{\partial W_t} p_{22}(V_A,W_t) = -\Im_{23}(V_A)$$

$$\frac{\partial}{\partial W_t}\ln p_{22}(V_A,W_t) = -\Im_{23}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t}\ln p_{22}(V_A,W_t)\,\partial W_t = \int_0^t -\Im_{23}(V_A)\,ds$$

$$\ln p_{22}(V_A,W_t) = [-\Im_{23}(V_A)t] \qquad p_{22}(V_A,W_t) = e^{-\Im_{23}(V_A)t} \qquad \textbf{(5.51)}$$

Solving for the derivative of equation (5.40) we obtained the transition probability of moving from stomach cancer state to Pharynx cancer state. The derivative of equation (5.40) is expressed as:

$$\frac{\partial}{\partial W_t} p_{23}(V_A,W_t) = p_{22}(V_A,W_t) * \Im_{23}(V_A) \qquad \frac{\partial}{\partial W_t} p_{23}(V_A,W_t) = e^{-\Im_{23}(V_A)t} * \Im_{23}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{23}(V_A,W_t)\,\partial W_t = \int_0^t e^{-\Im_{23}(V_A)s} * \Im_{23}(V_A)\,ds \qquad p_{23}(V_A,W_t) = \Im_{23}(V_A)\int_0^t e^{-\Im_{23}(V_A)s}\,ds$$

$$p_{23}(V_A,W_t) = \left[\frac{\Im_{23}(V_A) * e^{-\Im_{23}(V_A)s}}{-\Im_{23}(V_A)}\right]_0^t \qquad p_{23}(V_A,W_t) = 1 - e^{-\Im_{23}(V_A)t} \qquad \textbf{(5.52)}$$

Solving for the derivative of equation (5.42) we obtained the transition probability of remaining in Pharynx cancer state. The derivative of equation (5.42) is expressed as:

$$\frac{\partial}{\partial W_t} p_{33}(V_A,W_t) = p_{33}(V_A,W_t) * -\Im_{34}(V_A) \qquad \frac{1}{p_{33}(V_A,W_t)}\frac{\partial}{\partial W_t} p_{33}(V_A,W_t) = -\Im_{34}(V_A)$$

$$\frac{\partial}{\partial W_t}\ln p_{33}(V_A,W_t) = -\Im_{34}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t}\ln p_{33}(V_A,W_t)\,\partial W_t = \int_0^t -\Im_{34}(V_A)\,ds$$

$$\ln p_{33}(V_A,W_t) = [-\Im_{34}(V_A)s]^t_0 \qquad p_{33}(V_A,W_t) = e^{-\Im_{34}(V_A)t} \qquad \textbf{(5.53)}$$

Solving for the derivative of equation (5.44) we obtained the transition probability of transiting from pharynx state to colon cancer state. The derivative of equation (5.44) is expressed as:

$$\frac{\partial}{\partial W_t} p_{34}(V_A,W_t) = p_{33}(V_A,W_t) * \Im_{34}(V_A) \qquad \frac{\partial}{\partial W_t} p_{34}(V_A,W_t) = e^{-\Im_{34}(V_A)t} * \Im_{34}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{34}(V_A,W_t)\,\partial W_t = \int_0^t e^{-\Im_{34}(V_A)s} * \Im_{34}(V_A)\,ds \qquad p_{34}(V_A,W_t) = \Im_{34}(V_A)\int_0^t e^{-\Im_{34}(V_A)s}\,ds$$

$$p_{34}(V_A,W_t) = \left[\frac{\Im_{34}(V_A) * e^{-\Im_{34}(V_A)s}}{-\Im_{34}(V_A)}\right]_0^t \qquad p_{34}(V_A,W_t) = 1 - e^{-\Im_{34}(V_A)t} \qquad \textbf{(5.54)}$$

Solving for the derivative of equation (5.46) we obtained the transition probability of remaining in colon cancer state. The derivative of equation (5.46) is expressed as:

$$\frac{\partial}{\partial W_t} p_{44}(V_A,W_t) = p_{44}(V_A,W_t) * -\Im_{45}(V_A) \qquad \frac{1}{p_{44}(V_A,W_t)}\frac{\partial}{\partial W_t} p_{44}(V_A,W_t) = -\Im_{45}(V_A)$$

$$\frac{\partial}{\partial W_t}\ln p_{44}(V_A,W_t) = -\Im_{45}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t}\ln p_{44}(V_A,W_t)\,\partial W_t = \int_0^t -\Im_{45}(V_A)\,ds$$

$$\ln p_{44}(V_A,W_t) = [-\Im_{45}(V_A)t] \qquad p_{44}(V_A,W_t) = e^{-\Im_{45}(V_A)t} \qquad \textbf{(5.55)}$$

Solving for the derivative of equation (5.48) we obtained the transition probability of transiting from colon cancer state to dead state. The derivative of equation (5.48) is expressed as:

$$\frac{\partial}{\partial W_t} p_{45}(V_A, W_t) = p_{44}(V_A, W_t) * \Im_{45}(V_A) \qquad \frac{\partial}{\partial W_t} p_{45}(V_A, W_t) = e^{-\Im_{45}(V_A)t} * \Im_{45}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{45}(V_A, W_t)\partial W_t = \int_0^t e^{-\Im_{45}(V_A)s} * \Im_{45}(V_A) ds \qquad p_{45}(V_A, W_t) = \Im_{45}(V_A) \int_0^t e^{-\Im_{45}(V_A)s} ds$$

$$p_{45}(V_A, W_t) = \left[\frac{\Im_{45}(V_A) * e^{-\Im_{45}(V_A)s}}{-\Im_{45}(V_A)} \right]_0^t \qquad p_{45}(V_A, W_t) = 1 - e^{-\Im_{45}(V_A)t} \tag{5.56}$$

In this model state 5 is an absorbing state hence the transition probabilities are defined as: $p_{51}(V_A, W_t) = 0$, $p_{52}(V_A, W_t) = 0$, $p_{53}(V_A, W_t) = 0$, $p_{54}(V_A, W_t) = 0$, $p_{55}(V_A, W_t) = 1$
The transition probability matrix for this model is :

$$\begin{bmatrix} p_{11}(V_A, W_t) & p_{12}(V_A, W_t) & p_{13}(V_A, W_t) & p_{14}(V_A, W_t) & p_{15}(V_A, W_t) \\ p_{21}(V_A, W_t) & p_{22}(V_A, W_t) & p_{23}(V_A, W_t) & p_{24}(V_A, W_t) & p_{25}(V_A, W_t) \\ p_{31}(V_A, W_t) & p_{32}(V_A, W_t) & p_{33}(V_A, W_t) & p_{34}(V_A, W_t) & p_{35}(V_A, W_t) \\ p_{41}(V_A, W_t) & p_{42}(V_A, W_t) & p_{43}(V_A, W_t) & p_{44}(V_A, W_t) & p_{45}(V_A, W_t) \\ p_{51}(V_A, W_t) & p_{52}(V_A, W_t) & p_{53}(V_A, W_t) & p_{54}(V_A, W_t) & p_{55}(V_A, W_t) \end{bmatrix} =$$

$$\begin{bmatrix} e^{-\Im_{12}(V_A)t} & 1-e^{-\Im_{12}(V_A)t} & 0 & 0 & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1-e^{-\Im_{23}(V_A)t} & 0 & 0 \\ 0 & 0 & e^{-\Im_{34}(V_A)t} & 1-e^{-\Im_{34}(V_A)t} & 0 \\ 0 & 0 & 0 & e^{-\Im_{45}(V_A)t} & 1-e^{-\Im_{45}(V_A)t} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5.6 Six state cancer model

5.6.1 Introduction

In this section six state cancer Markov model which in cooperates Healthy state, Oesophagus cancer state, Stomach state ,Lung state, Kidney state and Dead state is derived. A case where patients can not move back to any state is considered. This leads to the assumption that a life cannot enter a state more than once. The patients who recover are assumed to have been censored from the study.

5.6.2 Six state Oesophagus cancer-Stomach-Lungs-Kidney model

Figure (5.4) represents the six-state model in which we systematically derive the respective Kolmogorov Forward Differential Equation.

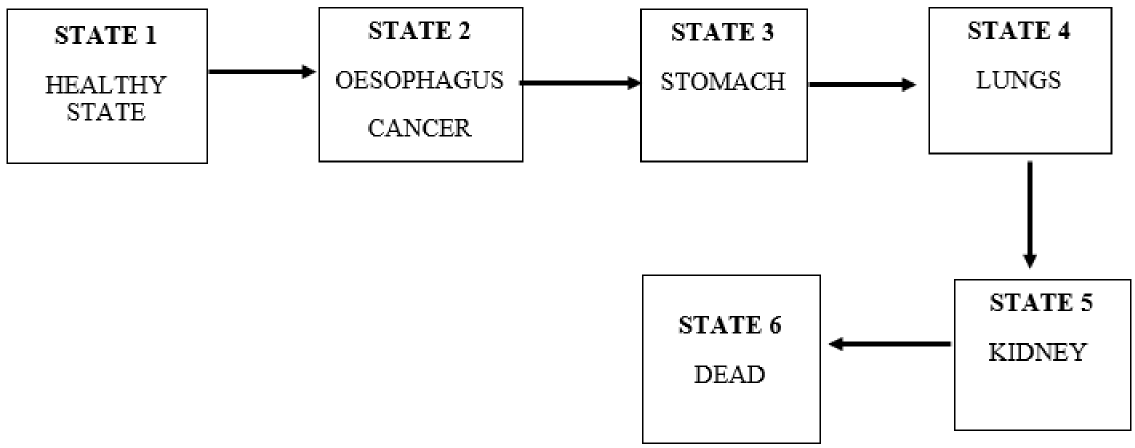


Figure 5.4. Oesophagus cancer-Stomach-Lungs-Kidney model

From the figure (5.4)above the transition probability matrix is expressed as:

$$\begin{bmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) & p_{15}(V_A,W_t) & p_{16}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) & p_{25}(V_A,W_t) & p_{26}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) & p_{35}(V_A,W_t) & p_{36}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) & p_{45}(V_A,W_t) & p_{46}(V_A,W_t) \\ p_{51}(V_A,W_t) & p_{52}(V_A,W_t) & p_{53}(V_A,W_t) & p_{54}(V_A,W_t) & p_{55}(V_A,W_t) & p_{56}(V_A,W_t) \\ p_{61}(V_A,W_t) & p_{62}(V_A,W_t) & p_{63}(V_A,W_t) & p_{64}(V_A,W_t) & p_{65}(V_A,W_t) & p_{66}(V_A,W_t) \end{bmatrix} =$$

$$\begin{bmatrix} a_4 & b_4 & 0 & 0 & 0 & 0 \\ 0 & c_4 & d_4 & 0 & 0 & 0 \\ 0 & 0 & e_4 & f_4 & 0 & 0 \\ 0 & 0 & 0 & g_4 & h_4 & 0 \\ 0 & 0 & 0 & 0 & k_4 & l_4 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Hence the transition probabilities of interest are; a_4, b_4, c_4, d_4, e_4, f_4, g_4, h_4, k_4 and l_4. Where a_4 represents transition probability of remaining in healthy state, b_4 represents the transition probability of moving from Healthy state to Oesophagus cancer state, c_4 represents the transition probability of remaining in Oesophagus cancer state, d_4 represents the transition probability of transiting from Oesophagus cancer state to Stomach cancer state, e_4 represents the transition probability of remaining in Stomach cancer state, f_4 is the transition probability of transiting from Stomach cancer state to Lung cancer state, g_4 represents the transition probability of remaining in Lung cancer state, h_4 is the transition probability of transiting from Lung cancer state to Kidney cancer state, k_4 represents the transition probability of remaining in Kidney cancer state and l_4 represents the transition probability of transiting from Kidney cancer state to Dead state.

Theorem 5.6.1 (Six state stomach-colon-liver-lung cancer model). *The transition probability matrix for this model is:*

$$\begin{pmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) & p_{15}(V_A,W_t) & p_{16}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) & p_{25}(V_A,W_t) & p_{26}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) & p_{35}(V_A,W_t) & p_{36}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) & p_{45}(V_A,W_t) & p_{46}(V_A,W_t) \\ p_{51}(V_A,W_t) & p_{52}(V_A,W_t) & p_{53}(V_A,W_t) & p_{54}(V_A,W_t) & p_{55}(V_A,W_t) & p_{56}(V_A,W_t) \\ p_{61}(V_A,W_t) & p_{62}(V_A,W_t) & p_{63}(V_A,W_t) & p_{64}(V_A,W_t) & p_{65}(V_A,W_t) & p_{66}(V_A,W_t) \end{pmatrix} =$$

$$\begin{pmatrix} e^{-\Im_{12}(V_A)t} & 1-e^{-\Im_{12}(V_A)t} & 0 & 0 & 0 & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1-e^{-\Im_{23}(V_A)t} & 0 & 0 & 0 \\ 0 & 0 & e^{-\Im_{34}(V_A)t} & 1-e^{-\Im_{34}(V_A)t} & 0 & 0 \\ 0 & 0 & 0 & e^{-\Im_{45}(V_A)t} & 1-e^{-\Im_{45}(V_A)t} & 0 \\ 0 & 0 & 0 & 0 & e^{-\Im_{56}(V_A)t} & 1-e^{-\Im_{56}(V_A)t} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Proof of theorem 5.6.1

Figure (5.4) represents the Six state Oesophagus-Stomach-Lung-Kidney cancer model. The transition intensities and probabilities are derived using Kolmogorov forward equations. Kolmogorov Forward Differential equation is expressed as:

$$p_{ij}(V_A,W_t+\kappa_d) = \sum_{k=1}^{n} p_{ik}(V_A,W_t)p_{kj}(W_t,W_t+\kappa_d)$$

In this model we consider the following: $(i,k,j) = 1,2,3,4,5,6$

When $i = 1$ and $j = 1$

$$\begin{aligned} p_{11}(V_A,W_t+\kappa_d) &= \sum_{k=1}^{6} p_{1k}(V_A,W_t)p_{k1}(W_t,W_t+\kappa_d) \\ &= p_{11}(V_A,W_t)p_{11}(W_t,W_t+\kappa_d) + p_{12}(V_A,W_t)p_{21}(W_t,W_t+\kappa_d) + \\ &\quad + p_{13}(V_A,W_t)p_{31}(W_t,W_t+\kappa_d) + p_{14}(V_A,W_t)p_{41}(W_t,W_t+\kappa_d) \\ &\quad + p_{15}(V_A,W_t)p_{51}(W_t,W_t+\kappa_d) + p_{16}(V_A,W_t)p_{61}(W_t,W_t+\kappa_d) \\ &= p_{11}(V_A,W_t)[1-p_{12}(W_t,W_t+\kappa_d)] \end{aligned} \quad \textbf{(5.57)}$$

Subtracting $p_{11}(V_A, W_t)$ from equation (5.57)

$$p_{11}(V_A, W_t + \kappa_d) - p_{11}(V_A, W_t) = p_{11}(V_A, W_t)[1 - p_{12}(W_t, W_t + \kappa_d)] - p_{11}(V_A, W_t)$$
$$= p_{11}(V_A, W_t)[-p_{12}(W_t, W_t + \kappa_d)]$$
$$\lim_{\kappa_d \to 0} \frac{p_{11}(V_A, W_t + \kappa_d) - p_{11}(V_A, W_t)}{\kappa_d} = p_{11}(V_A, W_t) \lim_{\kappa_d \to 0} \frac{-p_{12}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{11}(V_A, W_t) = p_{11}(V_A, W_t) * -\Im_{12}(V_A) \tag{5.58}$$

When $i = 1$ and $j = 2$

$$p_{12}(V_A, W_t + \kappa_d) = \sum_{k=1}^{6} p_{1k}(V_A, W_t) p_{k2}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)]$$
$$+ p_{13}(V_A, W_t) p_{32}(W_t, W_t + \kappa_d) + p_{14}(V_A, W_t) p_{42}(W_t, W_t + \kappa_d)$$
$$+ p_{15}(V_A, W_t) p_{52}(W_t, W_t + \kappa_d) + p_{16}(V_A, W_t) p_{62}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] \tag{5.59}$$

Subtracting $p_{12}(V_A, W_t)$ from equation (5.59)

$$p_{12}(V_A, W_t + \kappa_d) - p_{12}(V_A, W_t) = p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)$$
$$[1 - p_{23}(W_t, W_t + \kappa_d)] - p_{12}(V_A, W_t)$$
$$= p_{11}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{12}(V_A, W_t)$$
$$* -p_{23}(W_t, W_t + \kappa_d)$$
$$= p_{11}(V_A, W_t) * p_{12}(W_t, W_t + \kappa_d)$$

$$\lim_{\kappa_d \to 0} \frac{p_{12}(V_A, W_t + \kappa_d) - p_{12}(V_A, W_t)}{\kappa_d} = p_{11}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{12}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = p_{11}(V_A, W_t) * \Im_{12}(V_A) \tag{5.60}$$

When $i = 2$ and $j = 2$

$$p_{22}(V_A, W_t + \kappa_d) = \sum_{k=1}^{6} p_{2k}(V_A, W_t) p_{k2}(W_t, W_t + \kappa_d)$$
$$= p_{21}(V_A, W_t) p_{12}(W_t, W_t + \kappa_d) + p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)]$$
$$+ p_{23}(V_A, W_t) p_{32}(W_t, W_t + \kappa_d) + p_{24}(V_A, W_t) p_{42}(W_t, W_t + \kappa_d)$$
$$p_{25}(V_A, W_t) p_{52}(W_t, W_t + \kappa_d) + p_{26}(V_A, W_t) p_{62}(W_t, W_t + \kappa_d)$$
$$= p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] \tag{5.61}$$

Subtracting $p_{22}(V_A, W_t)$ from equation (5.61)

$$p_{22}(V_A, W_t + \kappa_d) - p_{22}(V_A, W_t) = p_{22}(V_A, W_t)[1 - p_{23}(W_t, W_t + \kappa_d)] - p_{22}(V_A, W_t)$$
$$= p_{22}(V_A, W_t) * -p_{23}(W_t, W_t + \kappa_d)$$
$$\lim_{\kappa_d \to 0} \frac{p_{22}(V_A, W_t + \kappa_d) - p_{22}(V_A, W_t)}{\kappa_d} = p_{22}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{23}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = p_{22}(V_A, W_t) * -\Im_{23}(V_A) \quad \textbf{(5.62)}$$

When $i = 2$ and $j = 3$

$$p_{23}(V_A, W_t + \kappa_d) = \sum_{k=1}^{6} p_{2k}(V_A, W_t) p_{k3}(W_t, W_t + \kappa_d)$$
$$= p_{21}(V_A, W_t) p_{13}(W_t, W_t + \kappa_d) + p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d)$$
$$+ p_{23}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d) + p_{24}(V_A, W_t) p_{43}(W_t, W_t + \kappa_d)$$
$$+ p_{25}(V_A, W_t) p_{53}(W_t, W_t + \kappa_d) + p_{26}(V_A, W_t) p_{63}(W_t, W_t + \kappa_d)$$
$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t) p_{33}(W_t, W_t + \kappa_d)$$
$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] \quad \textbf{(5.63)}$$

Subtracting $p_{23}(V_A, W_t)$ from equation (5.63)

$$p_{23}(V_A, W_t + \kappa_d) - p_{23}(V_A, W_t) = p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)$$
$$[1 - p_{34}(W_t, W_t + \kappa_d)] - p_{23}(V_A, W_t)$$
$$= p_{22}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d) + p_{23}(V_A, W_t)$$
$$[-p_{34}(W_t, W_t + \kappa_d)]$$
$$\lim_{\kappa_d \to 0} \frac{p_{23}(V_A, W_t + \kappa_d) - p_{23}(V_A, W_t)}{\kappa_d} = p_{22}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{23}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{23}(V_A, W_t) = p_{22}(V_A, W_t) * \Im_{23}(V_A) \quad \textbf{(5.64)}$$

When $i = 3$ and $j = 3$

$$p_{33}(V_A, W_t + \kappa_d) = \sum_{k=1}^{6} p_{3k}(V_A, W_t) p_{k3}(W_t, W_t + \kappa_d)$$
$$= p_{31}(V_A, W_t) p_{13}(W_t, W_t + \kappa_d) + p_{32}(V_A, W_t) p_{23}(W_t, W_t + \kappa_d)$$
$$+ p_{33}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] + p_{34}(V_A, W_t) p_{43}(W_t, W_t + \kappa_d)$$
$$+ p_{35}(V_A, W_t) p_{53}(W_t, W_t + \kappa_d) + p_{36}(V_A, W_t) p_{63}(W_t, W_t + \kappa_d)$$
$$= p_{33}(V_A, W_t)[1 - p_{34}(W_t, W_t + \kappa_d)] \quad \textbf{(5.65)}$$

Subtracting $p_{33}(V_A, W_t)$ from equation (5.65)

$$p_{33}(V_A, W_t+\kappa_d) - p_{33}(V_A, W_t) = p_{33}(V_A, W_t)[1 - p_{34}(W_t, W_t+\kappa_d)] - p_{33}(V_A, W_t)$$
$$= p_{33}(V_A, W_t)[-p_{34}(W_t, W_t+\kappa_d)]$$
$$\lim_{\kappa_d \to 0} \frac{p_{33}(V_A, W_t+\kappa_d) - p_{33}(V_A, W_t)}{\kappa_d} = p_{33}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{34}(W_t, W_t+\kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{33}(V_A, W_t) = p_{33}(V_A, W_t) * -\Im_{34}(V_A) \tag{5.66}$$

When $i = 3$ and $j = 4$

$$p_{34}(V_A, W_t+\kappa_d) = \sum_{k=1}^{6} p_{3k}(V_A, W_t) p_{k4}(W_t, W_t+\kappa_d)$$
$$+ p_{33}(V_A, W_t) p_{34}(W_t, W_t+\kappa_d) + p_{34}(V_A, W_t) p_{44}(W_t, W_t+\kappa_d)$$
$$+ p_{35}(V_A, W_t) p_{54}(W_t, W_t+\kappa_d) + p_{36}(V_A, W_t) p_{64}(W_t, W_t+\kappa_d)$$
$$= p_{33}(V_A, W_t) p_{34}(W_t, W_t+\kappa_d) + p_{34}(V_A, W_t)[1 - p_{45}(W_t, W_t+\kappa_d)] \tag{5.67}$$

Subtracting $p_{34}(V_A, W_t)$ from equation (5.67)

$$p_{34}(V_A, W_t+\kappa_d) - p_{34}(V_A, W_t) = p_{33}(V_A, W_t) p_{34}(W_t, W_t+\kappa_d) + p_{34}(V_A, W_t)$$
$$[1 - p_{45}(W_t, W_t+\kappa_d)] - p_{34}(V_A, W_t)$$
$$= p_{33}(V_A, W_t) p_{34}(W_t, W_t+\kappa_d) + p_{34}(V_A, W_t)$$
$$[-p_{45}(W_t, W_t+\kappa_d)]$$
$$= p_{33}(V_A, W_t) p_{34}(W_t, W_t+\kappa_d)$$
$$\lim_{\kappa_d \to 0} \frac{p_{34}(V_A, W_t+\kappa_d) - p_{34}(V_A, W_t)}{\kappa_d} = p_{33}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{34}(W_t, W_t+\kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{34}(V_A, W_t) = p_{33}(V_A, W_t) * \Im_{34}(V_A) \tag{5.68}$$

When $i = 4$ and $j = 4$

$$p_{44}(V_A, W_t+\kappa_d) = \sum_{k=1}^{6} p_{4k}(V_A, W_t) p_{k4}(W_t, W_t+\kappa_d)$$
$$= p_{41}(V_A, W_t) p_{14}(W_t, W_t+\kappa_d) + p_{42}(V_A, W_t) p_{24}(W_t, W_t+\kappa_d)$$
$$+ p_{43}(V_A, W_t) p_{34}(W_t, W_t+\kappa_d) + p_{44}(V_A, W_t)[1 - p_{45}(W_t, W_t+\kappa_d)]$$
$$+ p_{45}(V_A, W_t) p_{54}(W_t, W_t+\kappa_d) + p_{46}(V_A, W_t) p_{64}(W_t, W_t+\kappa_d)$$
$$= p_{44}(V_A, W_t)[1 - p_{45}(W_t, W_t+\kappa_d)] \tag{5.69}$$

Subtracting $p_{44}(V_A, W_t)$ from equation (5.69)

$$p_{44}(V_A, W_t + \kappa_d) = p_{44}(V_A, W_t)[1 - p_{45}(W_t, W_t + \kappa_d)]$$
$$= p_{44}(V_A, W_t)[-p_{45}(W_t, W_t + \kappa_d)]$$
$$\lim_{\kappa_d \to 0} \frac{p_{44}(V_A, W_t + \kappa_d) - p_{44}(V_A, W_t)}{\kappa_d} = p_{44}(V_A, W_t) * - \lim_{\kappa_d \to 0} \frac{p_{45}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{44}(V_A, W_t) = p_{44}(V_A, W_t) * -\Im_{45}(V_A) \quad \textbf{(5.70)}$$

When $i = 4$ and $j = 5$

$$p_{45}(V_A, W_t + \kappa_d) = \sum_{k=1}^{6} p_{4k}(V_A, W_t) p_{k5}(W_t, W_t + \kappa_d)$$
$$= p_{41}(V_A, W_t) p_{15}(W_t, W_t + \kappa_d) + p_{42}(V_A, W_t) p_{25}(W_t, W_t + \kappa_d)$$
$$+ p_{43}(V_A, W_t) p_{35}(W_t, W_t + \kappa_d) + p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d)$$
$$+ p_{45}(V_A, W_t)[1 - p_{56}(W_t, W_t + \kappa_d)] + p_{46}(V_A, W_t) p_{65}(W_t, W_t + \kappa_d)$$
$$= p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d) + p_{45}(V_A, W_t)[1 - p_{56}(W_t, W_t + \kappa_d)] \quad \textbf{(5.71)}$$

Subtracting $p_{45}(V_A, W_t)$ from equation (5.71)

$$p_{45}(V_A, W_t + \kappa_d) - p_{45}(V_A, W_t) = p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d) + p_{45}(V_A, W_t)$$
$$[1 - p_{56}(W_t, W_t + \kappa_d)] - p_{45}(V_A, W_t)$$
$$= p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d) + p_{45}(V_A, W_t)$$
$$[-p_{56}(W_t, W_t + \kappa_d)]$$
$$= p_{44}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d)$$
$$\lim_{\kappa_d \to 0} \frac{p_{45}(V_A, W_t + \kappa_d) - p_{45}(V_A, W_t)}{\kappa_d} = p_{44}(V_A, W_t) * \lim_{\kappa_d \to 0} \frac{p_{45}(W_t, W_t + \kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t} p_{45}(V_A, W_t) = p_{44}(V_A, W_t) * \Im_{45}(V_A) \quad \textbf{(5.72)}$$

When $i = 5$ and $j = 5$

$$p_{55}(V_A, W_t + \kappa_d) = \sum_{k=1}^{6} p_{5k}(V_A, W_t) p_{k5}(W_t, W_t + \kappa_d)$$
$$= p_{51}(V_A, W_t) p_{15}(W_t, W_t + \kappa_d) + p_{52}(V_A, W_t) p_{25}(W_t, W_t + \kappa_d)$$
$$+ p_{53}(V_A, W_t) p_{35}(W_t, W_t + \kappa_d) + p_{54}(V_A, W_t) p_{45}(W_t, W_t + \kappa_d)$$
$$+ p_{55}(V_A, W_t)[1 - p_{56}(W_t, W_t + \kappa_d)] + p_{56}(V_A, W_t) p_{65}(W_t, W_t + \kappa_d)$$
$$= p_{55}(V_A, W_t)[1 - p_{56}(W_t, W_t + \kappa_d)] \quad \textbf{(5.73)}$$

Subtracting $p_{55}(V_A,W_t)$ from equation (5.73)

$$p_{55}(V_A,W_t+\kappa_d) = p_{55}(V_A,W_t)[1-p_{56}(W_t,W_t+\kappa_d)] - p_{55}(V_A,W_t)$$
$$= p_{55}(V_A,W_t)[-p_{56}(W_t,W_t+\kappa_d)]$$
$$\lim_{\kappa_d\to 0}\frac{p_{55}(V_A,W_t+\kappa_d)-p_{55}(V_A,W_t)}{\kappa_d} = p_{55}(V_A,W_t) * -\lim_{\kappa_d\to 0}\frac{p_{56}(W_t,W_t+\kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t}p_{55}(V_A,W_t) = p_{55}(V_A,W_t) * -\Im_{56}(V_A) \qquad \textbf{(5.74)}$$

When $i=5$ and $j=6$

$$p_{56}(V_A,W_t+\kappa_d) = \sum_{k=1}^{6} p_{5k}(V_A,W_t)p_{k6}(W_t,W_t+\kappa_d)$$
$$= p_{51}(V_A,W_t)p_{16}(W_t,W_t+\kappa_d) + p_{52}(V_A,W_t)p_{26}(W_t,W_t+\kappa_d)$$
$$+ p_{53}(V_A,W_t)p_{36}(W_t,W_t+\kappa_d) + p_{54}(V_A,W_t)p_{46}(W_t,W_t+\kappa_d)$$
$$+ p_{55}(V_A,W_t)p_{56}(W_t,W_t+\kappa_d) + p_{56}(V_A,W_t)p_{66}(W_t,W_t+\kappa_d)$$
$$= p_{55}(V_A,W_t)p_{56}(W_t,W_t+\kappa_d) + p_{56}(V_A,W_t) \qquad \textbf{(5.75)}$$

Subtracting $p_{56}(V_A,W_t)$ from equation (5.75)

$$p_{56}(V_A,W_t+\kappa_d) - p_{56}(V_A,W_t) = p_{55}(V_A,W_t)p_{56}(W_t,W_t+\kappa_d) + p_{56}(V_A,W_t) - p_{56}(V_A,W_t)$$
$$= p_{55}(V_A,W_t)p_{56}(W_t,W_t+\kappa_d)$$
$$\lim_{\kappa_d\to 0}\frac{p_{56}(V_A,W_t+\kappa_d)-p_{56}(V_A,W_t)}{\kappa_d} = p_{55}(V_A,W_t) * \lim_{\kappa_d\to 0}\frac{p_{56}(W_t,W_t+\kappa_d)}{\kappa_d}$$
$$\frac{\partial}{\partial W_t}p_{56}(V_A,W_t) = p_{55}(V_A,W_t) * \Im_{56}(V_A) \qquad \textbf{(5.76)}$$

This model is build on the assumption that our observations are done within the interval time $(0,t)$. We solved individual derivatives to aNain the required transition probabilities.

Solving for the derivative of equation (5.58) we obtained the transition probability of moving or remaining in healthy state. The derivative of equation (5.58) is expressed as:

$$\frac{\partial}{\partial W_t}p_{11}(V_A,W_t) = p_{11}(V_A,W_t) * -\Im_{12}(V_A) \qquad \frac{1}{p_{11}(V_A,W_t)}\frac{\partial}{\partial W_t}p_{11}(V_A,W_t) = -\Im_{12}(V_A)$$
$$\frac{\partial}{\partial W_t}\ln p_{11}(V_A,W_t) = -\Im_{12}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t}\ln p_{11}(V_A,W_t)\partial W_t = \int_0^t -\Im_{12}(V_A)ds$$
$$\ln p_{11}(V_A,W_t) = [-\Im_{12}(V_A)t] \qquad p_{11}(V_A,W_t) = e^{-\Im_{12}(V_A)t} \qquad \textbf{(5.77)}$$

Solving for the derivative of equation (5.60) we obtained the transition probability of transiting from healthy state to oesophagus cancer state. The derivative of equation (5.60) is expressed as:

$$\frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = p_{11}(V_A, W_t) * \Im_{12}(V_A) \qquad \frac{\partial}{\partial W_t} p_{12}(V_A, W_t) = e^{-\Im_{12}(V_A)t} * \Im_{12}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{12}(V_A, W_t)\partial W_t = \int_0^t e^{-\Im_{12}(V_A)s} * \Im_{12}(V_A) ds \qquad p_{12}(V_A, W_t) = \Im_{12}(V_A)\int_0^t e^{-\Im_{12}(V_A)s} ds$$

$$p_{12}(V_A, W_t) = \left[\frac{\Im_{12}(V_A) * e^{-\Im_{12}(V_A)s}}{-\Im_{12}(V_A)}\right]_0^t \qquad p_{12}(V_A, W_t) = 1 - e^{-\Im_{12}(V_A)t} \tag{5.78}$$

Solving for the derivative of equation (5.62) we obtained the transition probability of remaining in oesophagus cancer state. The derivative of equation (5.62) is expressed as:

$$\frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = p_{22}(V_A, W_t) * -\Im_{23}(V_A) \qquad \frac{1}{p_{22}(V_A, W_t)}\frac{\partial}{\partial W_t} p_{22}(V_A, W_t) = -\Im_{23}(V_A)$$

$$\frac{\partial}{\partial W_t} \ln p_{22}(V_A, W_t) = -\Im_{23}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t} \ln p_{22}(V_A, W_t)\partial W_t = \int_0^t -\Im_{23}(V_A) ds$$

$$\ln p_{22}(V_A, W_t) = [-\Im_{23}(V_A)t] \qquad p_{22}(V_A, W_t) = e^{-\Im_{23}(V_A)t} \tag{5.79}$$

Solving for the derivative of equation (5.64) we obtained the transition probability of transiting from oesophagus cancer state to stomach cancer state. The derivative of equation (5.64) is expressed as:

$$\frac{\partial}{\partial W_t} p_{23}(V_A, W_t) = p_{22}(V_A, W_t) * \Im_{23}(V_A) \qquad \frac{\partial}{\partial W_t} p_{23}(V_A, W_t) = e^{-\Im_{23}(V_A)t} * \Im_{23}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{23}(V_A, W_t)\partial W_t = \int_0^t e^{-\Im_{23}(V_A)s} * \Im_{23}(V_A) ds \qquad p_{23}(V_A, W_t) = \Im_{23}(V_A)\int_0^t e^{-\Im_{23}(V_A)s} ds$$

$$p_{23}(V_A, W_t) = \left[\frac{\Im_{23}(V_A) * e^{-\Im_{23}(V_A)s}}{-\Im_{23}(V_A)}\right]_0^t \qquad p_{23}(V_A, W_t) = 1 - e^{-\Im_{23}(V_A)t} \tag{5.80}$$

Solving for the derivative of equation (5.66) we obtained the transition probability of remaining in stomach cancer state. The derivative of equation (5.66) is expressed as:

$$\frac{\partial}{\partial W_t} p_{33}(V_A, W_t) = p_{33}(V_A, W_t) * -\Im_{34}(V_A) \qquad \frac{1}{p_{33}(V_A, W_t)}\frac{\partial}{\partial W_t} p_{33}(V_A, W_t) = -\Im_{34}(V_A)$$

$$\frac{\partial}{\partial W_t} \ln p_{33}(V_A, W_t) = -\Im_{34}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t} \ln p_{33}(V_A, W_t)\partial W_t = \int_0^t -\Im_{34}(V_A) ds$$

$$\ln p_{33}(V_A, W_t) = [-\Im_{34}(V_A)t] \qquad p_{33}(V_A, W_t) = e^{-\Im_{34}(V_A)t} \tag{5.81}$$

Solving for the derivative of equation (5.68) we obtained the transition probability of transiting from stomach cancer state to lung cancer state. The derivative of equation (5.68) is expressed as:

$$\frac{\partial}{\partial W_t} p_{34}(V_A,W_t) = p_{33}(V_A,W_t) * \Im_{34}(V_A) \qquad \frac{\partial}{\partial W_t} p_{34}(V_A,W_t) = e^{-\Im_{34}(V_A)t} * \Im_{34}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{34}(V_A,W_t)\partial W_t = \int_0^t e^{-\Im_{34}(V_A)s} * \Im_{34}(V_A)ds \qquad p_{34}(V_A,W_t) = \Im_{34}(V_A)\int_0^t e^{-\Im_{34}(V_A)s}ds$$

$$p_{34}(V_A,W_t) = \left[\frac{\Im_{34}(V_A) * e^{-\Im_{34}(V_A)s}}{-\Im_{34}(V_A)}\right]_0^t \qquad p_{34}(V_A,W_t) = 1 - e^{-\Im_{34}(V_A)t} \tag{5.82}$$

Solving for the derivative of equation (5.70) we obtained the transition probability of remaining in lung cancer state. The derivative of equation (5.70) is expressed as:

$$\frac{\partial}{\partial W_t} p_{44}(V_A,W_t) = p_{44}(V_A,W_t) * -\Im_{45}(V_A) \qquad \frac{1}{p_{44}(V_A,W_t)}\frac{\partial}{\partial W_t} p_{44}(V_A,W_t) = -\Im_{45}(V_A)$$

$$\frac{\partial}{\partial W_t} \ln p_{44}(V_A,W_t) = -\Im_{45}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t} \ln p_{44}(V_A,W_t)\partial W_t = \int_0^t -\Im_{45}(V_A)ds$$

$$\ln p_{44}(V_A,W_t) = [-\Im_{45}(V_A)t] \qquad p_{44}(V_A,W_t) = e^{-\Im_{45}(V_A)t} \tag{5.83}$$

Solving for the derivative of equation (5.72) we obtained the transition probability of transiting from lung cancer state to kidney cancer state. The derivative of equation (5.72) is expressed as:

$$\frac{\partial}{\partial W_t} p_{45}(V_A,W_t) = p_{44}(V_A,W_t) * \Im_{45}(V_A) \qquad \frac{\partial}{\partial W_t} p_{45}(V_A,W_t) = e^{-\Im_{45}(V_A)t} * \Im_{45}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{45}(V_A,W_t)\partial W_t = \int_0^t e^{-\Im_{45}(V_A)s} * \Im_{45}(V_A)ds \qquad p_{45}(V_A,W_t) = \Im_{45}(V_A)\int_0^t e^{-\Im_{45}(V_A)s}ds$$

$$p_{45}(V_A,W_t) = \left[\frac{\Im_{45}(V_A) * e^{-\Im_{45}(V_A)s}}{-\Im_{45}(V_A)}\right]_0^t \qquad p_{45}(V_A,W_t) = 1 - e^{-\Im_{45}(V_A)t} \tag{5.84}$$

Solving for the derivative of equation (5.74) we obtained the transition probability of remaining in kidney cancer state. The derivative of equation (5.74) is expressed as:

$$\frac{\partial}{\partial W_t} p_{55}(V_A,W_t) = p_{55}(V_A,W_t) * -\Im_{56}(V_A) \qquad \frac{1}{p_{55}(V_A,W_t)}\frac{\partial}{\partial W_t} p_{55}(V_A,W_t) = -\Im_{56}(V_A)$$

$$\frac{\partial}{\partial W_t} \ln p_{55}(V_A,W_t) = -\Im_{56}(V_A) \qquad \int_0^t \frac{\partial}{\partial W_t} \ln p_{55}(V_A,W_t)\partial W_t = \int_0^t -\Im_{56}(V_A)ds$$

$$\ln p_{55}(V_A,W_t) = [-\Im_{56}(V_A)t] \qquad p_{55}(V_A,W_t) = e^{-\Im_{56}(V_A)t} \tag{5.85}$$

Solving for the derivative of equation (5.76) we obtained the transition probability of transiting from kidney cancer state to dead state. The derivative of equation (5.76) is expressed as:

$$\frac{\partial}{\partial W_t} p_{56}(V_A,W_t) = p_{55}(V_A,W_t) * \Im_{56}(V_A) \qquad \frac{\partial}{\partial W_t} p_{56}(V_A,W_t) = e^{-\Im_{56}(V_A)t} * \Im_{56}(V_A)$$

$$\int_0^t \frac{\partial}{\partial W_t} p_{56}(V_A,W_t)\partial W_t = \int_0^t e^{-\Im_{56}(V_A)s} * \Im_{56}(V_A)ds \qquad p_{56}(V_A,W_t) = \Im_{56}(V_A)\int_0^t e^{-\Im_{56}(V_A)s}ds$$

$$p_{56}(V_A,W_t) = \left[\frac{\Im_{56}(V_A) * e^{-\Im_{56}(V_A)s}}{-\Im_{56}(V_A)}\right]_0^t \qquad p_{56}(V_A,W_t) = 1 - e^{-\Im_{56}(V_A)t} \tag{5.86}$$

In this model state 6 is an absorbing state hence the transition probabilities are defined as:
$p_{61}(V_A,W_t) = 0$, $p_{62}(V_A,W_t) = 0$, $p_{63}(V_A,W_t) = 0$, $p_{64}(V_A,W_t) = 0$, $p_{65}(V_A,W_t) = 0$, $p_{66}(V_A,W_t) = 1$
The transition probability matrix for this model is :

$$\begin{pmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) & p_{15}(V_A,W_t) & p_{16}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) & p_{25}(V_A,W_t) & p_{26}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) & p_{35}(V_A,W_t) & p_{36}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) & p_{45}(V_A,W_t) & p_{46}(V_A,W_t) \\ p_{51}(V_A,W_t) & p_{52}(V_A,W_t) & p_{53}(V_A,W_t) & p_{54}(V_A,W_t) & p_{55}(V_A,W_t) & p_{56}(V_A,W_t) \\ p_{61}(V_A,W_t) & p_{62}(V_A,W_t) & p_{63}(V_A,W_t) & p_{64}(V_A,W_t) & p_{65}(V_A,W_t) & p_{66}(V_A,W_t) \end{pmatrix} =$$

$$\begin{pmatrix} e^{-\Im_{12}(V_A)t} & 1-e^{-\Im_{12}(V_A)t} & 0 & 0 & 0 & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1-e^{-\Im_{23}(V_A)t} & 0 & 0 & 0 \\ 0 & 0 & e^{-\Im_{34}(V_A)t} & 1-e^{-\Im_{34}(V_A)t} & 0 & 0 \\ 0 & 0 & 0 & e^{-\Im_{45}(V_A)t} & 1-e^{-\Im_{45}(V_A)t} & 0 \\ 0 & 0 & 0 & 0 & e^{-\Im_{56}(V_A)t} & 1-e^{-\Im_{56}(V_A)t} 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

5.7 Estimating transition and transition intensities

5.7.1 Introduction

Transition probability is the probabilities associated with various state changes and transition intensity is the the rate of change between states. Transition probabilities are derived from transition intensities or vice versa transition intensities can be derived from transition probabilities using Chapman Kolmogorov equations. This research considered estimation of transition probabilities and consequently estimating transition intensities. Product limit model was modified to estimate transition probabilities and the transition probabilities are used to calculate the transition intensities. Product limit model is based on the assumption that individuals under investigation transit at the same time hence only a single estimate is obtained.

5.7.2 Modified-product limit model

The product limit estimator is considered as the estimator of the survival function of lifetime data. Product limit is often considered in medical research in measuring the fraction of patients in a particular state for a certain amount of time after treatment. Product limit estimators can be considered in measuring the length of time people remain in a particular state after occurrence of a certain event. Product limit estimate is used to estimate survival estimate. Survival probabilities can also be estimated as:

$$S_x(t) = {}_t p_x = e^{-\mu t} \tag{5.87}$$

The probability of remaining in one state from the Markov models is defined as:

$$p_{ss}(V_A, W_t) = e^{-\Im_{sn}(V_A)t} \tag{5.88}$$

Let $\mu = \Im_{sn}$ hence combining equation (5.87) and equation (5.88) it becomes:

$$S_x(t) = e^{-\Im_{sn}(V_A)t} \tag{5.89}$$

Equation (5.89) can be estimated using Kaplan-Meier estimate.

Definition 5.7.1 (**Modified Kaplan-Meier estimate**). *Modified Kaplan-Meier estimate is expressed as:*

$$\hat{S}(t) = \prod_{t_j < t} (1 - \hat{\rho}_y) \tag{5.90}$$

where:
$\hat{S}(t)$ represents the probability of remaining in a particular state.
$\hat{\rho}_y$ is the hazard function which represents the number of affected cancer patients compared to the individuals under investigation.

Product limit estimate can be expressed in terms of transition probabilities as :

$$\hat{p}_{ss}(V_A, W_t) = \prod_{t_j < t} (1 - \hat{\rho}_y) \tag{5.91}$$

This research is based on the assumption that there is the risk under investigation occurs at one period hence equation (5.90) is modified to become:

$$\hat{S}(t) = (1 - \hat{\rho}_y) \tag{5.92}$$

Combining equation (5.91) and equation (5.92)

$$\hat{p}_{ss}(V_A, W_t) = (1 - \hat{\rho}_y) \tag{5.93}$$

where:
$1 - \hat{\rho}_y = \frac{n_y - m_y}{n_y}$, n_y represents the number of people at risk,
m_y is the number of individual affected by the risk under investigation.

Estimation of transition probabilities and transition intensities of Three stomach cancer model

The transition probabilities of interest in the three state stomach cancer model are expressed as:

$$p_{11}(V_A,W_t) = e^{-\Im_{12}(V_A)t} \qquad p_{12}(V_A,W_t) = 1 - e^{-\Im_{12}(V_A)t} \tag{5.94}$$

$$p_{22}(V_A,W_t) = e^{-\Im_{23}(V_A)t} \qquad p_{23}(V_A,W_t) = 1 - e^{-\Im_{23}(V_A)t} \tag{5.95}$$

Equation (5.94) can be further be expressed as:

$$p_{11}(V_A, W_t) = (1 - \hat{\rho}_y) \qquad = \frac{n_y - m_y}{n_y} \tag{5.96}$$

$$p_{12}(V_A, W_t) = 1 - (1 - \hat{\rho}_y) \qquad = 1 - \frac{n_y - m_y}{n_y} \tag{5.97}$$

The same methodology applied in equations (5.96) and (5.97) can be used for equation (5.95).

Estimation of transition probabilities and transition intensities of Four state breast-colon cancer model

The transition probabilities of interest in the four state stomach cancer model are expressed as:

$$p_{11}(V_A,W_t) = e^{-\Im_{12}(V_A)t} \qquad p_{12}(V_A,W_t) = 1 - e^{-\Im_{12}(V_A)t} \tag{5.98}$$

$$p_{22}(V_A,W_t) = e^{-\Im_{23}(V_A)t} \qquad p_{23}(V_A,W_t) = 1 - e^{-\Im_{23}(V_A)t} \tag{5.99}$$

$$p_{33}(V_A,W_t) = e^{-\Im_{34}(V_A)t} \qquad p_{34}(V_A,W_t) = 1 - e^{-\Im_{34}(V_A)t} \tag{5.100}$$

Equations (5.98), (5.99) and (5.100) can be further expressed as shown in equation (5.96) and (5.97).

Estimation of transition probabilities and transition intensities of Five state lung-stomach-colon cancer model

The transition probabilities of interest in the five state stomach cancer model are expressed as:

$$p_{11}(V_A,W_t) = e^{-\Im_{12}(V_A)t} \qquad p_{12}(V_A,W_t) = 1 - e^{-\Im_{12}(V_A)t} \tag{5.101}$$

$$p_{22}(V_A,W_t) = e^{-\Im_{23}(V_A)t} \qquad p_{23}(V_A,W_t) = 1 - e^{-\Im_{23}(V_A)t} \tag{5.102}$$

$$p_{33}(V_A,W_t) = e^{-\Im_{34}(V_A)t} \qquad p_{34}(V_A,W_t) = 1 - e^{-\Im_{34}(V_A)t} \tag{5.103}$$

$$p_{44}(V_A,W_t) = e^{-\Im_{45}(V_A)t} \qquad p_{45}(V_A,W_t) = 1 - e^{-\Im_{45}(V_A)t} \tag{5.104}$$

Equations (5.101), (5.102) , (5.103)and (5.104) can further be expressed as shown in equation (5.96) and (5.97).

Estimation of transition probabilities and transition intensities of Six state stomach-colon-liver-lung cancer model

The transition probabilities of interest in the six state stomach cancer model are expressed as:

$$p_{11}(V_A,W_t) = e^{-\Im_{12}(V_A)t} \qquad p_{12}(V_A,W_t) = 1 - e^{-\Im_{12}(V_A)t} \qquad \textbf{(5.105)}$$

$$p_{22}(V_A,W_t) = e^{-\Im_{23}(V_A)t} \qquad p_{23}(V_A,W_t) = 1 - e^{-\Im_{23}(V_A)t} \qquad \textbf{(5.106)}$$

$$p_{33}(V_A,W_t) = e^{-\Im_{34}(V_A)t} \qquad p_{34}(V_A,W_t) = 1 - e^{-\Im_{34}(V_A)t} \qquad \textbf{(5.107)}$$

$$p_{44}(V_A,W_t) = e^{-\Im_{45}(V_A)t} \qquad p_{45}(V_A,W_t) = 1 - e^{-\Im_{45}(V_A)t} \qquad \textbf{(5.108)}$$

$$p_{55}(V_A,W_t) = e^{-\Im_{56}(V_A)t} \qquad p_{56}(V_A,W_t) = 1 - e^{-\Im_{56}(V_A)t} \qquad \textbf{(5.109)}$$

Equations (5.105), (5.106), (5.107), (5.108) and (5.109) can further be expressed as shown in equation (5.96) and (5.97).

5.8 Applicability in Discrete phase type distributions

Discrete phase type distributions are build on a Markov chain where one state is an absorbing state while the other state are transient. The matrices derived in this chapter should satisfy that condition for them to be applicable in phase type distributions. For three state model the multi state model obtained is:

$$\left[\begin{array}{cc|c} e^{-\Im_{12}(V_A)t} & 1-e^{-\Im_{12}(V_A)t} & 0 \\ 0 & e^{-\Im_{23}(V_A)t} & 1-e^{-\Im_{23}(V_A)t} \\ \hline 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{cc} Z & y \\ 0 & 1 \end{array}\right]$$

This is the representation required for discrete phase type distribution which represents distribution of the time to absorption of a Markov chain.

5.9 Chapter summary

The main objective of this chapter was to develop multi-state models of secondary cancer cases to be applied as the matrix parameters of the phase type models. Multi-state models of four selected secondary cancer cases are developed for: Leukemia cancer model, Liver cancer-Colon model, Stomach cancer-Pharynx-Colon model and Oesophagus cancer-Stomach-Lung-Kidney model. The transition intensities and stationary probabilities for each model are developed.

6 SEVERITY DISTRIBUTIONS

6.1 Introduction

Severity distributions are applied in modeling of claim amounts. Continuous distributions and discrete distribution are considered in this research. Severity probabilities are applied in Panjer recursive model and Discrete Fourier in order to estimate aggregate loss probabilities. Continuous distribution that are considered are:

(i) Weibull distribution

(ii) Generalized Pareto distribution

(iii) Pareto distribution

Discrete distributions considered in this research are:

i One parameter Poisson Lindley distribution

ii Two parameter Poisson Lindley distribution

6.2 Continuous distributions

Continuous distribution are discretized in order to be applied in Panjer recursive formula and Discrete Fourier Transform which requires only discrete distributions. Method of rounding is to discretized the continuous distributions.

6.2.1 Weibull distribution

A continuous random variable is said to follow Weibull distribution if its pdf is given by:

$$f(x) = \begin{cases} \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} & ; x > 0; \alpha, \beta > 0 \\ 0 & \text{otherwise} \end{cases}$$

where α and β are the parameters.

Cumulative distribution function

The cumulative distribution of Weibull distribution can be derived from its probability density function.

$$F(x) = p(x \leq x) = \int_0^x f(x)dx = \int_0^x \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} dx$$

Let $\left(\frac{x}{\alpha}\right)^{\beta} = y$ hence $\frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}dx = dy$

$$F(x) = \int_0^{\left(\frac{x}{\alpha}\right)^{\beta}} e^{-y}dy = -\left[e^{-y}\right]_0^{\left(\frac{x}{\alpha}\right)^{\beta}} = -\left[e^{-\left(\frac{x}{\alpha}\right)^{\beta}} - 1\right]$$

Therefore:

$$F(x) = \begin{cases} 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}} & ; x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Mean and Variance of Weibull distribution

The r^{th} moment about the origin are:

$$\mu_r^1 = E[X^r] = \int_0^{\infty} x^r f(x)dx = \int_0^{\infty} x^r \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} dx$$

Let $\left(\frac{x}{\alpha}\right)^{\beta} = y$, $x = \alpha y^{\frac{1}{\beta}}$ and $\frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}dx = dy$

$$\mu_r^1 = \int_0^{\infty} \alpha^r y^{\frac{r}{\beta}} e^{-y}dy = \alpha^r \int_0^{\infty} y^{\frac{r}{\beta}} e^{-y}dy$$

But $\Gamma n = \int_0^{\infty} e^{-x}x^{n-1}dx$

Therefore:

$$\mu_r^1 = \alpha^r \Gamma\left(1 + \frac{r}{\beta}\right) \quad \textbf{(6.1)}$$

When $r = 1$ we get the expectation of Weibull distribution hence it is expressed as:

$$E(x) = \alpha \Gamma\left(1 + \frac{1}{\beta}\right) \quad \textbf{(6.2)}$$

When $r = 2$ we get $E(x^2)$ which is expressed as:

$$\mu_2^1 = \alpha^2 \Gamma\left(1 + \frac{2}{\beta}\right) \quad \textbf{(6.3)}$$

Variance of Weibull distribution can be calculated using equation (6.2) and equation (6.3). The variance is expressed as:

$$Var(x) = \mu_2^1 - [\mu_1^1]^2 = \alpha^2\Gamma\left(1+\frac{2}{\beta}\right) - \alpha^2\left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2$$
$$= \alpha^2\left(\Gamma\left(1+\frac{2}{\beta}\right) - \left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2\right) \quad \textbf{(6.4)}$$

Estimation of parameters using Method of Moments

Expectation is calculated using the formula:

$$E(x) = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Variance is calculated using the formula that:

$$Var(x) = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \frac{1}{N}\sum_{i=1}^{N} x_i$$

Combining equation (6.2) and (6.4) results to:

$$Var(x) = \alpha^2\Gamma\left(1+\frac{2}{\beta}\right) - [E(x)]^2 \quad \textbf{(6.5)}$$

Equation (6.2) can be rearranged as:

$$\alpha = \frac{E(x)}{\Gamma(1+\frac{1}{\beta})}$$

Hence equation (6.5) becomes:

$$Var(x) = \frac{[E(x)]^2}{[\Gamma(1+\frac{1}{\beta})]^2} * \Gamma\left(1+\frac{2}{\beta}\right) - [E(x)]^2 \qquad \frac{Var(x) - [E(x)]^2}{[E(x)]^2} = \frac{\Gamma(1+\frac{2}{\beta})}{[\Gamma(1+\frac{1}{\beta})]^2}$$

6.2.2 Generalized Pareto distribution

Generalized Pareto distribution is constructed by mixing Gamma distribution and Gamma distribution. Let the conditional Gamma distribution be expressed as:

$$f(x|\beta) = \frac{\beta^{\alpha}}{\Gamma\alpha} x^{\alpha-1} e^{-\beta x} \quad \textbf{(6.6)}$$

and the Gamma distribution as:

$$g(\beta) = \frac{\lambda^{\gamma}}{\Gamma(\gamma)} \beta^{\gamma-1} e^{-\beta\lambda} \quad \textbf{(6.7)}$$

Mixing the distribution in equation (6.6) and distribution in equation (6.7) results to:

$$f(x) = \int_0^{\infty} f(x|\beta) g(\beta)\, d\beta \qquad = x^{\alpha-1} \int_0^{\infty} \frac{\beta^{\alpha}}{\Gamma(\alpha)} e^{-\beta x} * \frac{\lambda^{\gamma}}{\Gamma(\gamma)} \beta^{\gamma-1} e^{-\beta\lambda}\, d\beta$$

Hence:

$$f(x) = \frac{x^{\alpha-1}\lambda^{\gamma}}{\Gamma(\gamma)\Gamma(\alpha)} \int_0^{\infty} e^{-\beta(x+\lambda)} \beta^{\alpha+\gamma+1}\, d\beta \quad \textbf{(6.8)}$$

Let

$$y = \beta(x+\lambda) \qquad \beta = \frac{y}{x+\lambda}$$

The derivative if β is:

$$d\beta = \frac{dy}{x+\lambda}$$

Replacing the derivative of β in equation (6.8) results to:

$$f(x) = \frac{x^{\alpha-1}\lambda^{\gamma}}{\Gamma(\gamma)\Gamma(\alpha)} \int_0^{\infty} \frac{y^{\alpha+\gamma-1}}{(x+\lambda)^{\alpha+\gamma-1}} e^{-y} \frac{dy}{x+\lambda} \qquad = \frac{x^{\alpha-1}\lambda^{\gamma}\Gamma(\alpha+\gamma)}{\Gamma(\gamma)\Gamma\alpha(x+\lambda)^{\alpha+\gamma}}$$

Hence:

$$f(x) = \frac{x^{\alpha-1}}{\beta(\alpha,\gamma)(x+\lambda)^{\alpha+\gamma}} dx \quad \textbf{(6.9)}$$

Mean and Variance of Generalized Pareto

The first moment of Generalized Pareto distribution is expressed as:

$$E(x) = \frac{\gamma\lambda}{\alpha-1} \quad \textbf{(6.10)}$$

The second moment of Generalized Pareto distribution is expressed as:

$$Var(x) = \frac{\gamma\lambda^{2}(\gamma+\alpha-1)}{(\alpha-1)^{2}(\alpha-2)} \quad \textbf{(6.11)}$$

Estimation of parameters using Method of Moments

Generalized Pareto distribution can also be expressed as;

$$f(x) = \frac{1}{e}\left[1 - \frac{d(x-f)}{e}\right]^{\frac{1}{d}-1} \qquad \text{assuming that} \quad X = e(1 - exp(-dY))/d.$$

and a is the shape parameter, b is the scale parameter and c is the location parameter. The moment estimators of Generalized Pareto distribution is expressed as:

$$E(x) = f + \frac{e}{1+d} \qquad Var(x) = \frac{e^2}{(1+d)^2(1+2d)} \qquad Skew(x) = \frac{2(1-d)(1+2d)^{0.5}}{1+3d}$$

The value of d can be calculated from the skewness formula,hence e and f can be estimated as follows:

$$e = S.D(1+d)(1+2d)^{0.5} \qquad f = \bar{x} - \frac{e}{e+d}$$

The cumulative distribution function is expressed as:

$$F_X = 1 - \left[1 - \frac{d(x-f)}{e}\right]^{\frac{1}{d}}$$

6.2.3 Pareto distribution

Pareto distribution is constructed by mixing Gamma distribution and exponential distribution. The exponential distribution is expressed as:

$$f(x|\theta) = \theta e^{-\theta x} \qquad x > 0, \theta > 0 \tag{6.12}$$

The gamma distribution is expressed as:

$$g(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\theta} \theta^{\alpha-1} \tag{6.13}$$

The Pareto distribution is derived as by mixing equation (6.12) and equation (6.13).

$$f(x) = \int_0^\infty \theta e^{-\theta x} g(\theta)\, d\theta = \int_0^\infty \theta e^{-\theta x} \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\theta} \theta^{\alpha-1} d\theta = \frac{\beta^\alpha}{\Gamma(\alpha-1)} \frac{\Gamma\alpha}{(x+\beta)^\alpha}$$

Hence :

$$f(x) = \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}} \tag{6.14}$$

Mean and Variance of Pareto distribution

The first and second of Pareto distribution is expressed respectively as:

$$E(x) = \frac{\beta}{\alpha - 1} \quad \alpha > 1 \qquad Var(x) = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)} \quad \alpha > 2 \tag{6.15}$$

The k^{th} moment of and cdf of Pareto distribution and is expressed respectively as:

$$E[x^k] = \frac{\beta^k k!}{(\alpha-1)\ldots(\alpha-k)} \qquad F(x) = 1 - \left(\frac{\beta}{\beta + x}\right)^{\alpha} \tag{6.16}$$

Estimation of parameters using Method of moments

Estimation of values of α and β using method of moments. The value of α is estimated as follows:

$$E(x) = \frac{\beta}{\alpha - 1} \quad \alpha > 1 \quad \text{and} \quad E(x^2) = \frac{2\beta^2}{(\alpha-1)(\alpha-2)}$$

β can be estimated by using these two expressions.

$$\frac{E(x^2)}{[E(x)]^2} = \frac{\frac{2\beta^2}{(\alpha-1)(\alpha-2)}}{\frac{\beta^2}{(\alpha-1)^2}} = \frac{2\beta^2}{(\alpha-1)(\alpha-2)}\frac{(\alpha-1)^2}{\beta^2}(\alpha-1) = \frac{2(\alpha-1)}{(\alpha-2)} \tag{6.17}$$

6.3 Discretization of claim severity

Claim amount distribution should be a scale distribution. A scale distribution is a distribution that if a random variable from that distribution is multiplied by a positive constant the resulting distribution belonging to the same family of distribution of the original random variable. Distributions which satisfy the condition of scale distribution are Pareto distribution,Weibull distribution Exponential distribution among others. Calculation of aggregate loss distribution using Panjer recursive model and Discrete Fourier Transform requires a discrete distribution hence method of rounding, also known as method of mass dispersal is used in discretizing the continuous distributions.

6.3.1 Method of rounding or method of mass dispersal

This method is used to convert the severity distribution to an equispaced arithmetic distribution. Severity distribution are usually mostly continuous distributions. A span which fit the data according to how large the data is chosen for discretization. This method hugely relies on the probability one-half span on either side of jh and places it at jh.The following formulas are used to discretize

the continuous severity.
The initial probability is calculated using the formula:

$$f_0 = p_r\left(j < \frac{h}{2}\right) = F_j\left(\frac{h}{2}\right)$$

The subsequent probabilities are calculated using the formula:

$$f_x = p_r\left(xh - \frac{h}{2} \le j \le xh + \frac{h}{2}\right) = F_j\left(xh + \frac{h}{2}\right) - F_j\left(xh - \frac{h}{2}\right) \quad x = 1, 2, 3, \ldots$$

The discretazation process is halved at some point when most of the probability has been accounted for. The process is halved when all the f_x adds up to 1 to ensure that the discretization process leads to a probability density function. At this point it is expressed as :

$$f_m = 1 - F_J\left(mh - \frac{h}{2}\right)$$

6.4 Discrete distributions

Discrete distributions can be directly applied in Panjer recursive model hence there is no need for discretization.

6.4.1 One parameter Poisson Lindley distribution

The probability mass function of one parameter Poison Lindley is expressed as:

$$f(x) = \frac{\theta^2(x + \theta + 2)}{(\theta + 1)^{x+3}} \quad x = 0, 1, 2, \ldots \tag{6.18}$$

Mean and variance of one parameter Poisson Lindley distribution

The r^{th} factorial moments about the origin is expressed as:

$$\mu_r^1 = E[E(X^r/\lambda)]$$

where $X^r = X(X-1)(X-2)(X-3)\ldots(X-k+1)$

$$\mu_r^1 = \int_0^\infty p(x|\lambda)f(\lambda;\theta)d\lambda = \int_0^\infty \lambda^r\left[\sum_{x=r}^{\infty}\frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right]\frac{\theta^2}{\theta+1}(1+\lambda)e^{-\theta\lambda}d\lambda$$

Let $x + k$ replace x hence:

$$\mu_r^1 = \int_0^\infty \lambda^r\left[\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^{x}}{x!}\right]\frac{\theta^2}{\theta+1}(1+\lambda)e^{-\theta\lambda}d\lambda \tag{6.19}$$

where $\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}$ is the pdf of Poisson distribution hence it is equal to 1, hence equation (6.19) becomes:

$$\mu_r^1 = \int_0^\infty \lambda^r \frac{\theta^2}{\theta+1}(1+\lambda)e^{-\theta\lambda}\,d\lambda = \frac{r!(\theta+r+1)}{\theta^r(\theta+1)} \quad r = 1,2,3,... \tag{6.20}$$

when $r = 1$ equation (6.20) is the expectation of one parameter Poisson Lindley distribution which is expressed as:

$$\mu_1' = \frac{1!(\theta+r+1)}{\theta^r(\theta+1)} = \frac{\theta+2}{\theta(\theta+1)} \tag{6.21}$$

The variance of one parameter Poisson Lindley distribution is expressed as:

$$\mu_2' = \frac{2!(\theta+r+1)}{\theta^r(\theta+1)} = \frac{\theta+2}{\theta(\theta+1)} + \frac{2(\theta+3)}{\theta^2(\theta+1)} \tag{6.22}$$

Estimation of parameter of using method of moments

θ can be estimated using method of moments as :

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \hat{\theta} = \frac{-(\bar{X}-1)+\sqrt{(\bar{X}-1)^2+8\bar{X}}}{2\bar{X}} \tag{6.23}$$

6.4.2 Two parameter Poisson Lindley

Two Poisson Lindley distribution is expressed as:

$$f(x) = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)} \frac{\theta^2}{\alpha\theta+1}(\alpha+x)e^{-\theta\lambda}\,d\lambda = \frac{\theta^2}{(\theta+1)^{x+2}}\left[1+\frac{\alpha+x}{\alpha\theta+1}\right] \quad \theta > 0; \alpha\theta > -1 \tag{6.24}$$

The r^{th} moments about the origin of two- parameter Poisson Lindley distribution is expressed as:

$$\mu_r' = E\left[E\left(\frac{X^r}{\lambda}\right)\right] \tag{6.25}$$

Combining equation (6.24) and equation (6.25) it becomes:

$$\mu_r' = \int_0^\infty \left[\sum_{x=0}^{\infty} x^r \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)}\right] \frac{\theta^2}{\alpha\theta+1}(\alpha+x)e^{-\theta\lambda}\,d\lambda \tag{6.26}$$

when $r = 1$ we can evaluate mean of two parameter Poisson Lindley distribution:

$$\mu_1' = \int_0^\infty \left[\sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)}\right] \frac{\theta^2}{\alpha\theta+1}(\alpha+x)e^{-\theta\lambda}\,d\lambda = \frac{\alpha\theta+2}{\theta(\alpha\theta+1)} \tag{6.27}$$

when $r = 2$ we can evaluate variance of two parameter Poisson Lindley distribution:

$$\mu_2' = \int_0^\infty \left[\sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)} \right] \frac{\theta^2}{\alpha\theta+1}(\alpha+x)\, e^{-\theta\lambda}\, d\lambda \qquad = \frac{\alpha\theta+2}{\theta(\alpha\theta+1)} + \frac{2(\alpha\theta+3)}{\theta^2(\alpha\theta+1)} \tag{6.28}$$

Parameter estimation using method of moments

Combining equation (6.27) and equation (6.28) it becomes:

$$\frac{\mu_2' - \mu_1'}{[\mu_1']^2} = \frac{2(\alpha\theta+3)(\alpha\theta+1)}{(\alpha\theta+2)^2} \qquad \frac{\mu_2' - \mu_1'}{[\mu_1']^2} = \frac{2(b+3)(b+1)}{(b+2)^2} \tag{6.29}$$

Let $\frac{\mu_2' - \mu_1'}{[\mu_1']^2}$ be k , hence equation (6.29) becomes;

$$k = \frac{2(b+3)(b+1)}{(b+2)^2}$$

$$0 = (2-k)b^2 + (8-4k)b + (6-4k) \tag{6.30}$$

The value of k can be calculated hence equation (6.30) can be solved. Substituting $b = \alpha\theta$ in equation (6.27) becomes;

$$\bar{X} = \frac{b+2}{\theta(b+1)} \tag{6.31}$$

and thus:

$$\theta = \frac{b+2}{(b+1)\bar{X}} \tag{6.32}$$

The estimate of α can be expressed as :

$$\hat{\alpha} = \frac{b(b+1)\bar{X}}{b+2} \tag{6.33}$$

6.5 Chapter summary

The main objective of this chapter was to highlight the severity distributions used in modeling claim severity data and methods of parameter estimation. This chapter highlights continuous and discrete distributions considered in modeling claim severity data as well as their parameter estimations.

7 ESTIMATION OF AGGREGATE LOSSES USING PH PANJER RECURSION AND DFT

7.1 Introduction

Phase type Panjer recursive models for class $(a, b, 0)$ and class $(a, b, 1)$ are used to estimate the aggregate loss probabilities using Panjer recursive model and one, two parameter Poisson Lindley, Zero-truncated one and two parameter Poisson Lindley distributions are also considered in estimation of aggregate loss probabilities using Discrete Fourier Transform. The distributions considered for frequency distributions are:

(i) Panjer class $(a, b, 0)$ distributions.

(ii) Panjer class $(a, b, 1)$ distributions.

(iii) Phase type one parameter Poisson Lindley distribution.

(iv) Phase type two parameter Poisson Lindley distribution.

(v) Phase type Zero-truncated one parameter Poisson Lindley distribution.

(vi) Phase type Zero-truncated two parameter Poisson Lindley distribution.

This research considers both continuous and discrete distributions in modeling of claim severity. The continuous distributions considered are:

(i) Pareto distribution.

(ii) Generalized Pareto distribution.

(iii) Weibull distribution

The discrete distributions are mixture of Poisson distribution and Lindley distribution hence they are not discretized and they are:

(i) One parameter Poisson-Lindley distribution.

(ii) Two parameter Poisson-Lindley distribution.

7.2 Phase type recursive model for class $(a, b, 0)$

The initial condition P_0 of phase type Panjer recursive model is estimated as well as the matrices A and B. Phase type Panjer recursive model for class $(a, b, 0)$ is expressed as:

$$\vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(A + B\frac{x}{j}\right)[I - Ay(0)]^{-1} \tag{7.1}$$

where $\vec{Z}(0) = \vec{Y}P_0$

7.2.1 Phase type Poisson distribution

Phase type Poisson distribution is expressed as:

$$P_n = \frac{e^{-\Lambda}\Lambda^{x}}{x!} \tag{7.2}$$

where Λ is $M * M$ matrix. The value of Λ is estimated using Markov chains. The initial condition of phase type Poisson distribution is expressed as:

$$P_0 = e^{-\Lambda} \tag{7.3}$$

The matrices A and B for PH Poisson are:

$$A = 0 \qquad B = \Lambda \tag{7.4}$$

Combining equation (7.1) and (7.4) it becomes:

$$\vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(0 + \Lambda\frac{x}{j}\right)[I - 0y(0)]^{-1} \qquad \vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(\Lambda\frac{x}{j}\right) \tag{7.5}$$

7.2.2 Phase type Negative Binomial distribution

Phase type Negative Binomial distribution is expressed as:

$$P_n = \binom{n+\alpha-1}{n}[I-Q]^{\alpha}Q^{n} \tag{7.6}$$

where Q is $M * M$ matrix. The value of Q is estimated using Markov chains. The initial condition of phase type Negative Binomial distribution is expressed as:

$$P_0 = [I-Q]^{\alpha} \tag{7.7}$$

The matrices A and B for PH Negative Binomial can be shown as:

$$A = Q \qquad B = (\alpha - 1)Q \tag{7.8}$$

Combining equation (7.1) and (7.8) it becomes:

$$\vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(Q + [(\alpha-1)Q]\frac{x}{j}\right)[I - Qy(0)]^{-1} \tag{7.9}$$

7.2.3 Phase type Geometric distribution

Phase type Geometric distribution is expressed as:

$$P_n = [I - P]^n P \tag{7.10}$$

where P is $M * M$ matrix.The value of P is estimated using Markov chains. The initial condition of phase type Geometric distribution is expressed as:

$$P_0 = P \tag{7.11}$$

The matrices A and B for PH Geometric are:

$$A = [I - P] \qquad B = 0 \tag{7.12}$$

Combining equation (7.1) and (7.12) it becomes:

$$\vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)[I-P]\left(I - [I-P]y(0)\right)^{-1} \tag{7.13}$$

7.2.4 Phase type Binomial distribution

Phase type Binomial distribution is expressed as:

$$P_n = \binom{\alpha}{n} P^n [I-P]^{\alpha-n} \tag{7.14}$$

where P is $M * M$ matrix.The value of P is estimated using Markov chains. The initial condition of phase type Binomial distribution is expressed as:

$$P_0 = [I - P]^{\alpha} \tag{7.15}$$

The matrices A and B for PH Binomial are:

$$A = -P[I-P]^{-1} \qquad B = (\alpha+1)P[I-P]^{-1} \tag{7.16}$$

Combining equation (7.1) and (7.16) it becomes:

$$\vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(-P[I-P]^{-1} + (\alpha+1)P[I-P]^{-1}\frac{x}{j}\right)\left(I + P[I-P]^{-1}y(0)\right)^{-1} \tag{7.17}$$

7.3 Phase type recursive model for class $(a, b, 1)$

The initial conditions P_0 and P_1 of phase type Panjer recursive model are estimated as well as the matrices A and B. Phase type Panjer recursive model for class $(a, b, 1)$ is expressed as:

$$\vec{Z}(j) = [P_1 - (A+B)P_0]y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(A + B\frac{x}{j}\right)[I - Ay(0)]^{-1} \tag{7.18}$$

7.3.1 Phase type Zero Truncated Poisson distribution

Phase type Zero Truncated Poisson distribution is expressed as:

$$P_n = \Lambda^n (e^{-\Lambda} - I)^{-1} \tag{7.19}$$

where Λ is $M * M$ matrix. The value of Λ is estimated using Markov chains. The initial conditions of phase type Zero Truncated Poisson distribution is expressed as:

$$P_0 = e^{-\Lambda} \tag{7.20}$$

Hence P_1 is expressed as:

$$P^T_k = P_k[I - P_0]^{-1} \qquad k = 1, 2, 3, \ldots \qquad P^T_1 = P_1[I - P_0]^{-1} \tag{7.21}$$

The matrices A and B for PH Zero truncated Poisson are:

$$A = 0 \qquad B = \Lambda \tag{7.22}$$

Combining equation (7.18), (7.20), (7.21) and (7.22) it becomes:

$$\vec{Z}(j) = [P_1 - (0 + \Lambda P_0]y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(0 + \Lambda\frac{x}{j}\right)[I - 0y(0)]^{-1}$$

$$\vec{Z}(j) = [P_1 - \Lambda P_0]y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(\Lambda\frac{x}{j}\right) \tag{7.23}$$

7.3.2 Phase type Zero Truncated Binomial distribution

Phase type Zero Truncated Binomial distribution is expressed as:

$$P_n = \binom{\alpha}{n} P^n Q^{\alpha-n} \left(I - Q^{\alpha}\right)^{-1} \tag{7.24}$$

where P is $M * M$ matrix.The value of P is estimated using Markov chains. The initial conditions of phase type Zero Truncated Binomial distribution is expressed as:

$$P_0 = [I - P]^{\alpha} \qquad \textbf{(7.25)}$$

Hence P_1 is expressed as:

$$P_k^T = P_k[I - P_0]^{-1} \qquad k = 1, 2, 3, \ldots \qquad P_1^T = P_1[I - P_0]^{-1} \qquad \textbf{(7.26)}$$

The matrices A and B for Zero truncated Binomial are:

$$A = -P[I - P]^{-1} \qquad B = (\alpha + 1)P[I - P]^{-1} \qquad \textbf{(7.27)}$$

Combining equation (7.18), (7.25), (7.26) and (7.27) it becomes:

$$\vec{Z}(j) = \left[P_1 - \left[-P[I-P]^{-1} + (\alpha+1)P[I-P]^{-1} \right] P_0 \right] y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x) \left[-P[I-P]^{-1} + (\alpha+1)P[I-P]^{-1} \frac{x}{j} \right] \left[I + P[I-P]^{-1} y(0) \right]^{-1} \qquad \textbf{(7.28)}$$

7.3.3 Phase type Zero Truncated Geometric distribution

Phase type Zero Truncated Geometric distribution is expressed as:

$$P_n = P[I - P]^{n-1} \qquad \textbf{(7.29)}$$

where P is $M * M$ matrix.The value of P is estimated using Markov chains. The initial conditions of phase type Zero Truncated Geometric distribution is expressed as:

$$P_0 = P \qquad \textbf{(7.30)}$$

Hence P_1 is expressed as:

$$P_k^T = P_k[I - P_0]^{-1} \qquad k = 1, 2, 3, \ldots \qquad P_1^T = P_1[I - P_0]^{-1} \qquad \textbf{(7.31)}$$

The matrices A and B for PH Zero truanted Geometric are:

$$A = [I - P] \qquad B = 0 \qquad \textbf{(7.32)}$$

Combining equation (7.18), (7.30), (7.31) and (7.32) it becomes:

$$\vec{Z}(j) = [P_1 - [(I-P) + 0]P_0]y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x) \left[[I-P] + 0\frac{x}{j} \right] \left[I - [I-P]y(0) \right]^{-1}$$

$$\vec{Z}(j) = \Upsilon[P_1 - (I-P)P_0]y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x)[I-P][I-[I-P]y(0)]^{-1} \qquad \textbf{(7.33)}$$

7.4 Discrete Fourier Transform

DFT is a computed function hence they are used to compute compound distribution. Compound distribution can be evaluated when both claim count and claim distribution are known. DFT is said to be a way of representing function in terms of a point value representation (that is a very specific value representation). DFT converts a finite sequence of equally spaced samples of a function into a same length of equally spaced samples which is a complex valued function of frequency. The DFT uses the characteristic function to convert the sequences. The probability generating function of the frequency distribution is illustrated as:

$$\Phi(n) = \sum_{k=0}^{\infty} n^k P_k \tag{7.34}$$

The characteristic function of the claim severity density is defined as:

$$\Phi(t) = \int_{-\infty}^{\infty} f(x)e^{itx}\, dx \tag{7.35}$$

Let the characteristic function of the claim count distribution be:

$$\phi_X(t) = \int_0^{\infty} e^{itx} dF_X(x) = \sum_{k=1}^{n} d_k \frac{e^{itx_k} - e^{itx_{k-1}}}{it} + f_{n+1}e^{itx_n} \tag{7.36}$$

The characteristic has real and imaginary parts which can be separated by using Euler's formula :

$$e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

The real part is:

$$y(t) = Re[\phi_x(t)] = \frac{1}{t}\sum_{k=1}^{n} d_k[\sin(tx_k) - \sin(tx_{k-1})] + f_{n+1}\cos(tx_n)$$

The imaginary part is:

$$z(t) = lm[\phi_X(t)] = \frac{1}{t}\sum_{k=1}^{n} d_k[\cos(tx_{k-1}) - \cos(tx_k)] + f_{n+1}\sin(tx_n)$$

The pgf of aggregate loss distribution can be expressed as :

$$P_S(t) = P_N[P_X(t)]. \tag{7.37}$$

The characteristic function of the aggregate loss distribution can be expressed as:

$$\phi_S(t) = E[e^{iS_t}] = P_N[\phi_x(t)]. \tag{7.38}$$

The characteristic function of the aggregate loss distribution in equation (7.38) can be obtained using DFT as:

$$\phi_S(t) = P_N[y(t) + iZ(t)] \quad \textbf{(7.39)}$$

Equation (7.39) has complex numbers hence it can be re wriNen as :

$$\phi_S(t) = r(t)e^{i\theta(t)} \quad \textbf{(7.40)}$$

DFT algorithm of aggregate losses is computed using DFT of claim count and DFT of claim amount separately. Let X_z be the claim amount distribution of the claim data. For any continuous function X_k the Fourier transform is;

$$X_k = \int_{-\infty}^{x} X_z e^{itx} dx$$

The inverse of this Fourier transform can be wriNen as ;

$$X_z = \frac{1}{2\Pi} \int_{-\infty}^{\infty} X_k e^{itx} dt$$

When X_k is a probability function of a discrete distribution then the Fourier transform is known as a DFT. The characteristic function of the claim severity density function will be derived from discretizing the the Fourier transform. DFT transforms a sequence of N complex numbers $\{X_n\} = X_0, X_1, \ldots, X_{Y-1}$ into another sequence of complex number $\{X_k\} = X_0, X_1, \ldots, X_{Y-1}$. This sequence of complex numbers can either be the severity probabilities or the frequency probabilities.

In the case of discretized severity probabilities its DFT is expressed as:

$$X_k = \sum_{z=0}^{Y-1} X(z) e^{\frac{-i2\Pi kz}{Y}} \qquad k = 0, 1, 2, \ldots, Y-1 \quad \textbf{(7.41)}$$

Expression (7.41) is very complex hence Euler's formula is used to reduce its complexity. Euler formula is expressed in two ways.

Case I: for $-i$

$$e^{-ix} = \cos x - i \sin x \quad \textbf{(7.42)}$$

Case II: for i

$$e^{ix} = \cos x + i \sin x \quad \textbf{(7.43)}$$

$\cos x$ can be expressed as:

$$\cos x = Re(e^{ix}) = \frac{e^{ix} + e^{-ix}}{2}$$

$\sin x$ can be expressed as:

$$\sin x = Im\left(e^{ix}\right) = \frac{e^{ix} - e^{ix}}{2i}$$

Applying Eulers formula equation (7.41) it becomes:

$$X(k) = \sum_{z=0}^{Y-1} X(z) \left[\cos\left(\frac{2\Pi kz}{Y}\right) - i\sin\left(\frac{2\Pi kz}{Y}\right)\right] \quad \textbf{(7.44)}$$

Let $\cos(\frac{2\Pi}{Y}) - i\sin(\frac{2\Pi}{Y})$ be expressed as W_Z.
This simplifies equation (7.44) to :

$$X(k) = \sum_{z=0}^{Y-1} X(z) W_Y^{kz} \quad \textbf{(7.45)}$$

This is the DFT of the discretized claim amount probabilities. The same procedure is followed in estimation of DFT of frequency distribution only that the frequency probabilities do not require discretization because the frequency distribution is already a discrete distribution.

7.4.1 No wrap convolution

The distribution of sum of two random variables is given by the convolution of the respective distributions. The convolution of vectors is done by multiplying the two vectors.
let $a = (a_0, a_1, a_2, \ldots, a_{n-1})$ and $b = (b_0, b_1, b_2, \ldots, b_{n-1})$ be vectors of the same length n. The discrete convolution of these vectors, $c = a * b$ is vector of length n defined by:

$$c_i = \sum_{j=0}^{n-1} a_j b_{i-j} \qquad 0 \leq i \leq n-1$$

The convolution required in DFT is no wrap convolution.No wrap convolution of vector a and b is defined to have the following components:

$$c_i = \sum_{j=0}^{i} a_j b_{i-j}$$

No wrap convolution is done by taking one vector,reversing it and placing it so that its first element is directly below the first element of the other vector. The vector are then successively shifted together,elements in the same column multiplied and the products added.This is repeated until the vectors are completely aligned.

7.4.2 Estimation of aggregate loss distribution using DFT

The frequency and severity probabilities are lengthened with equal number of zero's as its elements. The matrix W_N^{kz} is multiplied with the claim count and claim amount probabilities to compute the DFT of the claim count and claim amount probabilities respectively. The DFT of claim count probabilities is multiplied with the DFT of the claim amount probabilities. The resulting vector from the product of the DFT of frequency and DFT of severity is multiplied with the matrix W_Y^{kz} to get the DFT of DFT of claim count and DFT of claim amount probabilities. The values without the complex i are singled out and each value divided by the number of elements in the vector of claim count or claim amount distribution(i.e The vectors should have the same number of elements).The resulting probabilities are arranged in reverse except for the first probability.The values that correspond to claim count and claim amount distribution original values before the padding they become the aggregate loss distribution.

7.5 Chapter summary

The main objective of this chapter was to develop aggregate loss models to be applied in modeling of secondary cancer cases. Aggregate loss models using PH Panjer class $(a, b, 0)$ and PH Panjer class $(a, b, 1)$ are developed for each specific distribution. Discrete Fourier Transform used in estimating aggregate losses is also developed.

8 DATA ANALYSIS AND RESULTS

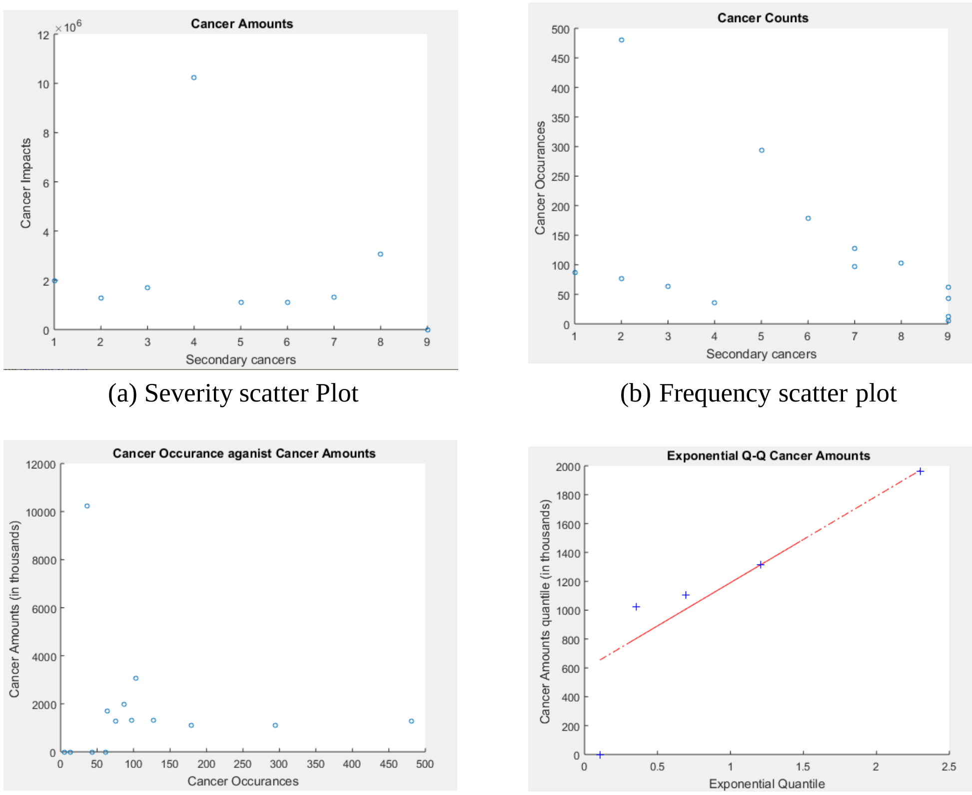
8.1 Introduction

In this chapter transition probabilities of three state model, four state model, five state model and six state model are estimated and their transition intensities. The parameters of severity and frequency distributions are estimated and consequently used in estimation of aggregate loss probabilities. Phase type Panjer recursive model and Discrete Fourier Transform are used to estimate the aggregate loss probabilities. The data taken into consideration was secondary cancer data obtained from a health facility in Kenya. The models developed in this research can be employed in different data sets all over the world. The frequency data is quantified the number of cancer patients affected by secondary cancers while severity data is quantified by the amount used to treat the patients. In this research we considered a period of five years, enabling us to capture the transitions. Descriptive statistics for secondary cancer claim count and secondary cancer claim amount are represented in table (8.1). The least number of secondary cancer claim count is five which represents the least number of deaths related to one of the secondary cancers investigated. The least secondary claim amount number is zero which represent amount related to secondary cancer deaths as it does not aNract any treatment cost. 481 was the highest number of secondary cancer claim count while $10,235,000$ was the highest for secondary cancer claim amount. The expectations of this data varied depending on the different states considered as illustrated in table (8.1). The data considered in this research for three state model, four state model, five state model and six state model has its expectation less than its median for all variable except for severity data for six state, indicating that the data is left-skewed.

Statistics	Cancer Occurrences	Cancer Amounts
N	850	850
N for States		
3	103	103
4	179	179
5	481	481
6	87	87
Total	850	21,735,857
Mean		
3	41	339,723
4	35	268,985
5	57	409,702
6	11	1,684,999
Standard Deviation		
3	21	960,881
4	44	505,720
5	91	582,156
6	21	3,118,754
Minimum	5	0
1st Quantile	48	1,104,000
Median	82	1,316,865
3rd Quantile	121	1,959,330
Maximum	481	10,235,000
Inter Quantile Range	73	855,330

Table 8.1. Descriptive statistics of occurrence and severity of secondary cancer data

Figure 8.1 illustrates exploratory data analysis for secondary cancer data. Figure 8.1 shows scaNer-plots for claim count data and claim amount data which shows claim counts and claim amounts have not seriously violated the independence assumption. The exponential Q-Q plot shows deviation from the straight line on the lower part indicating that cancer claim amount data is left skewed.



(a) Severity scatter Plot (b) Frequency scatter plot

(c) Comparison scatter plot (d) Exponential

Figure 8.1. Scater Plots and Q-Q Plot of Cancer data

8.2 Estimation of transition probabilities and transition intensity

The concept of survival probabilities is be used to estimate the transition probabilities.

8.2.1 Three state multi-state model

The probability of remaining in state 1 , $p_{11}(V_A, W_t) = e^{-\mathfrak{I}_{12}(V_A)t}$, is equal to the probability of survival, implying that the individuals did not transit to the next cancer state. The probability of remaining in a particular state can be calculated using the formula:

$$p_{jj}(V_A, W_t) = (1 - \hat{\rho}_y) \qquad = \frac{n_y - m_y}{n_y} \tag{8.1}$$

The probability of moving from one state to the next can be estimated using the formula:

$$p_{jk}(V_A, W_t) = 1 - (1 - \hat{\rho}_y) \qquad = 1 - \frac{n_y - m_y}{n_y} \tag{8.2}$$

The transition probabilities for the three state model estimated using Matlab software are:

$$\begin{pmatrix} p_{11}(V_A, W_t) & p_{12}(V_A, W_t) & p_{13}(V_A, W_t) \\ p_{21}(V_A, W_t) & p_{22}(V_A, W_t) & p_{23}(V_A, W_t) \\ p_{31}(V_A, W_t) & p_{32}(V_A, W_t) & p_{33}(V_A, W_t) \end{pmatrix} = \begin{pmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{pmatrix}$$

The transition intensity from state 1 to 2 and state 2 to 3 are calculated as:

$$p_{jj}(V_A,W_t) = e^{-\Im_{j(j+1)}(V_A)t} \qquad \Im_{12}(V_A) = 0.0258 \qquad \Im_{23}(V_A) = 0.1842$$

8.2.2 Four state multi-state model

The probability of remaining in a particular state can be calculated using equation (8.1) and the probability of moving from one state to the next can be estimated using equation (8.2). The transition probabilities for the four state model estimated using Matlab software are:

$$\begin{bmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) \end{bmatrix} = \begin{bmatrix} 0.7894 & 0.2106 & 0 & 0 \\ 0 & 0.2905 & 0.7095 & 0 \\ 0 & 0 & 0.8976 & 0.1024 \\ 0 & 0 & 0 & 1.0000 \end{bmatrix}$$

The transition intensity from state 1 to 2 , state 2 to 3 and state 3 to 4 are calculated as:

$$p_{jj}(V_A,W_t) = e^{-\Im_{j(j+1)}(V_A)t} \qquad \Im_{12}(V_A) = 0.0473 \quad \Im_{23}(V_A) = 0.2472 \quad \Im_{34}(V_A) = 0.0216$$

8.2.3 Five state multi-state model

The probability of remaining in a particular state can be calculated using equation (8.1) and the probability of moving from one state to the next can be estimated using equation (8.2). The transition probabilities for the five state model estimated using Matlab software are:

$$\begin{bmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) & p_{15}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) & p_{25}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) & p_{35}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) & p_{45}(V_A,W_t) \\ p_{51}(V_A,W_t) & p_{52}(V_A,W_t) & p_{53}(V_A,W_t) & p_{54}(V_A,W_t) & p_{55}(V_A,W_t) \end{bmatrix}$$

$$= \begin{bmatrix} 0.4341 & 0.5659 & 0 & 0 & 0 \\ 0 & 0.3888 & 0.6112 & 0 & 0 \\ 0 & 0 & 0.6701 & 0.3299 & 0 \\ 0 & 0 & 0 & 0.5567 & 0.4433 \\ 0 & 0 & 0 & 0 & 1.0000 \end{bmatrix}$$

The transition intensity from state 1 to 2 , state 2 to 3 , state 3 to 4 and state 4 to 5 are calculated as:

$$p_{jj}(V_A,W_t) = e^{-\Im_{j(j+1)}(V_A)t} \qquad \Im_{12}(V_A) = 0.1667 \qquad \Im_{23}(V_A) = 0.1889$$

$$\Im_{34}(V_A) = 0.0801 \qquad \Im_{45}(V_A) = 0.1171$$

8.2.4 Six state multi-state model

The probability of remaining in a particular state can be calculated using equation (8.1) and the probability of moving from one state to the next can be estimated using equation (8.2). The transition probabilities for the six state model estimated using Matlab software are:

$$\begin{bmatrix} p_{11}(V_A,W_t) & p_{12}(V_A,W_t) & p_{13}(V_A,W_t) & p_{14}(V_A,W_t) & p_{15}(V_A,W_t) & p_{16}(V_A,W_t) \\ p_{21}(V_A,W_t) & p_{22}(V_A,W_t) & p_{23}(V_A,W_t) & p_{24}(V_A,W_t) & p_{25}(V_A,W_t) & p_{26}(V_A,W_t) \\ p_{31}(V_A,W_t) & p_{32}(V_A,W_t) & p_{33}(V_A,W_t) & p_{34}(V_A,W_t) & p_{35}(V_A,W_t) & p_{36}(V_A,W_t) \\ p_{41}(V_A,W_t) & p_{42}(V_A,W_t) & p_{43}(V_A,W_t) & p_{44}(V_A,W_t) & p_{45}(V_A,W_t) & p_{46}(V_A,W_t) \\ p_{51}(V_A,W_t) & p_{52}(V_A,W_t) & p_{53}(V_A,W_t) & p_{54}(V_A,W_t) & p_{55}(V_A,W_t) & p_{56}(V_A,W_t) \\ p_{61}(V_A,W_t) & p_{62}(V_A,W_t) & p_{63}(V_A,W_t) & p_{64}(V_A,W_t) & p_{65}(V_A,W_t) & p_{66}(V_A,W_t) \end{bmatrix}$$

$$= \begin{bmatrix} 0.8976 & 0.1024 & 0 & 0 & 0 & 0 \\ 0 & 0.1264 & 0.8736 & 0 & 0 & 0 \\ 0 & 0 & 0.1579 & 0.8421 & 0 & 0 \\ 0 & 0 & 0 & 0.4375 & 0.5625 & 0 \\ 0 & 0 & 0 & 0 & 0.8611 & 0.1389 \\ 0 & 0 & 0 & 0 & 0 & 1.0000 \end{bmatrix}$$

The transition intensity from state 1 to 2 , state 2 to 3 , state 3 to 4 , state 4 to 5 and state 5 to 6 are calculated as:

$$p_{jj}(V_A,W_t) = e^{-\Im_{j(j+1)}(V_A)t} \qquad \Im_{12}(V_A) = 0.0216 \qquad \Im_{23}(V_A) = 0.4137$$

$$\Im_{34}(V_A) = 0.3691 \qquad \Im_{45}(V_A) = 0.1653 \qquad \Im_{56}(V_A) = 0.0299$$

8.3 Estimation of stationary probabilities for the multi-state models

Ɏ represents the stationary probabilities of the multi-state Markov models hence it needs to be estimated for each multi-state model.

8.3.1 Three state Leukemia model

Let the three state Leukemia model be expressed as A. Stationary probabilities are evaluated using the formula:

$$\pi = \pi A \tag{8.3}$$

where π is row vector of the stationary distribution. Because of the nature of the matrix A it is not possible to evaluate the stationary probabilities directly hence the stationary probabilities are evaluated at a point k where there is the least error values between the stationary probabilities when fiNed in the formula $\pi = \pi A$. The initial probability π_0 is expressed as:

$$\pi_0 = [1, 0, 0]$$

This is because every individual considered in this research was alive and healthy at the beginning of this investigation. Consequently the stationary probabilities are evaluated as:

$$\pi_1 = \pi_0 A \qquad \pi_2 = \pi_1 A \qquad \pi_2 = \pi_0 A^2 \qquad \pi_k = \pi_0 A^k \tag{8.4}$$

The stationary probabilities for three state Leukemia model is evaluated using equation (8.4) as:

$$\pi_{11} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix}^{11}$$

$$\begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \\ 0 & 0.0000 & 1.0000 \\ 0 & 0 & 1.0000 \end{bmatrix}$$

Replacing the matrix $[0.2122 \;\; 0.0535 \;\; 0.7343]$ in equation (8.3) it becomes :

$$\begin{bmatrix} 0.2108 & 0.0529 & 0.7363 \end{bmatrix} = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} * \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix}$$

This has an error of 0.02 which is negligible, hence the stationary probabilities for three state Leukemia model are:

$$\pi_{3\,state} = [0.2122 \;\; 0.0535 \;\; 0.7343] \tag{8.5}$$

The same methodology is used for three state model, four state model five state model and six state model resulting to:

States	π_1	π_2	π_3	π_4	π_5	π_6
4 state	0.0000	0.0000	0.0001	0.9999		
5 state	0.0039	0.0014	0.0053	0.0066	0.9827	
6 state	0.0000	0.0000	0.0000	0.0000	0.0001	0.9999

Table 8.2. Stationary probabilities

8.4 Estimation of frequency probabilities and its moments for class $(a, b, 0)$

This Section estimates frequency probabilities for phase type Poisson, Negative Binomial, Binomial and Geometric distribution.

8.4.1 Probabilities of Phase type Poisson distribution

Phase type Poisson distribution is expressed as:

$$p_n = \Upsilon \frac{e^{-\Lambda}\Lambda^n}{n!} 1^T \quad \textbf{(8.6)}$$

Λ is a $M * M$ matrix which is evaluated as $3 * 3$ matrix, $4 * 4$ matrix, $5 * 5$ matrix, and $6 * 6$ matrix in this research. The values of Λ and Υ have already been estimated hence equation (8.6) can be evaluated. The values of Λ and Υ are:

$$\Lambda = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \quad \textbf{(8.7)}$$

Replacing values of three, four, five and six state model in equation (8.6) the probabilities respectively are:

$$p_3 = 0.0245 \qquad p_4 = 0.0184 \qquad p_5 = 0.0147 \qquad p_6 = 0.0123 \quad \textbf{(8.8)}$$

Estimating moments of Phase type Poisson distribution

The mean and variance of phase type Poisson distribution are evaluated using the formulas;

$$E(\hat{N}) = \Upsilon\Lambda 1^T \qquad Var(\hat{N}) = \Upsilon\Lambda 1^T \quad \textbf{(8.9)}$$

Replacing equation (8.9) with values of stationary probabilities and matrix parameter for three, four, five and six state model the mean and variance respectively are:

$$E(\hat{N}) = 1.0001 \qquad E(\hat{N}) = 1.0000 \qquad E(\hat{N}) = 1.0001 \qquad E(\hat{N}) = 1.0000$$
$$Var(\hat{N}) = 1.001 \qquad Var(\hat{N}) = 1 \qquad Var(\hat{N}) = 1.001 \qquad Var(\hat{N}) = 1$$

8.4.2 Probabilities of Phase type Negative Binomial distribution

Phase type Negative Binomial distribution is expressed as:

$$p_n = \Upsilon \binom{n + \alpha - 1}{n} [I - Q]^{\alpha} Q^n 1^T \qquad n = 0, 1, 2, \ldots \quad \textbf{(8.10)}$$

Q is a $(M * M)$ matrix which is evaluated as $(3 * 3)$ matrix,$(4 * 4)$ matrix,$(5 * 5)$ matrix, and $(6 * 6)$ matrix in this research. The value of α for three, four, five and six state model can be evaluate using the equation of the mean as the value of Q has already been evaluated. The value of α for three, four, five and six state model can be evaluated respectively as shown in equation (8.11) as:

$$\alpha = \frac{E(\hat{N})}{\Upsilon Q[I-Q]^{-1}\mathbb{1}^T} \qquad \alpha_3 = 4.2025 \quad \alpha_4 = 0.64439 \quad \alpha_5 = 0.39506 \quad \alpha_6 = 0.281395 \qquad \mathbf{(8.11)}$$

The values of Q and Υ have already been estimated hence equation (8.10) can be evaluated. The values of Q and Υ are:

$$Q = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \qquad \mathbf{(8.12)}$$

Replacing values of three, four, five and six state model in equation (8.10) the probabilities respectively are:

$$p_3 = 2.1865 * e^{-19} \qquad p_4 = 2.5411 * e^{-22} \qquad p_5 = 0 \qquad p_6 = 1.1294 * e^{-22}$$

Estimating moments of Phase Negative Binomial distribution

The mean and variance of phase type Negative Binomial can be evaluated using the formula;

$$E(\hat{N}) = \alpha\Upsilon Q[I-Q]^{-1}\mathbb{1}^T \qquad Var(\hat{N}) = \alpha\Upsilon\{Q^2[I-Q]^{-2} + Q[I-Q]^{-1}\}\mathbb{1}^T \qquad \mathbf{(8.13)}$$

Replacing equation (8.13) with values of stationary probabilities, values of α for each state and matrix parameter for three, four, five and six state model the mean and variance respectively are:

$$E(\hat{N}) = 6.8116 \qquad E(\hat{N}) = 5.0592e^{-04} \qquad E(\hat{N}) = 0.0106 \qquad E(\hat{N}) = 1.5577e^{-04}$$
$$Var(\hat{N}) = 11.0033 \qquad Var(\hat{N}) = 8.9616e^{-04} \qquad Var(\hat{N}) = 0.0135 \qquad Var(\hat{N}) = 2.6274e^{-04}$$

8.4.3 Probabilities of Phase type Binomial distribution

Phase type Binomial distribution is expressed as:

$$p_n = \Upsilon \binom{\alpha}{n} [P]^n Q^{\alpha-n}\mathbb{1}^T \qquad n = 0, 1, 2, \ldots \qquad \mathbf{(8.14)}$$

P is a $(M * M)$ matrix which is evaluated as $(3 * 3)$ matrix,$(4 * 4)$ matrix,$(5 * 5)$ matrix, and $(6 * 6)$ matrix in this research. The value of α can be evaluate using the equation of the mean as the value of P has already been evaluated.

$$\alpha = \frac{E(\hat{N})}{\Upsilon P 1^T} = 470 \qquad \textbf{(8.15)}$$

The values of P and Υ have already been estimated hence equation (8.14) can be evaluated. The values of P and Υ are:

$$P = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \qquad \textbf{(8.16)}$$

Replacing values of three, four, five and six state model and α in equation (8.14) the probabilities respectively are:

$$p_3 = 2.2457e^{-102} \qquad p_4 = 0 \qquad p_5 = 1.5888e^{-102} \qquad p_6 = 0 \qquad \textbf{(8.17)}$$

Estimating moments of Phase Binomial distribution

The mean and variance can be evaluated using the formula;

$$E(\hat{N}) = \alpha \Upsilon P 1^T \qquad Var(\hat{N}) = \alpha \Upsilon \{PQ\} 1^T \qquad \textbf{(8.18)}$$

Replacing equation (8.18) with values of stationary probabilities, values of α for each state and matrix parameter for three state model, four state model, five state model and six state model the mean and variance respectively are:

$$E(\hat{N}) = 471.0471 \qquad E(\hat{N}) = 471 \qquad E(\hat{N}) = 471.0471 \qquad E(\hat{N}) = 471$$
$$Var(\hat{N}) = 6.5364e^{-15} \qquad Var(\hat{N}) = 0 \qquad Var(\hat{N}) = 0 \qquad Var(\hat{N}) = 1.5958e^{-18}$$

8.4.4 Probabilities of Phase type Geometric distribution

Phase type Geometric distribution is expressed as:

$$p_n = \Upsilon [I - P]^n P 1^T \qquad n = 0, 1, 2, \ldots \qquad \textbf{(8.19)}$$

P is a $(M * M)$ matrix is evaluated as $(3 * 3)$ matrix, $(4 * 4)$ matrix, $(5 * 5)$ matrix, and $(6 * 6)$ matrix in this research. The values of P and Υ have already been estimated hence equation (8.19) can be evaluated. The values of P and Υ are:

$$P = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \qquad \textbf{(8.20)}$$

Replacing values of three state model, four state model, five state model and six state model in equation (8.19) the probabilities respectively are:

$$p_3 = 3.4694e^{-18} \qquad p_4 = 1.4476e^{-24} \qquad p_5 = 0 \qquad p_6 = 3.3087e^{-24} \qquad \textbf{(8.21)}$$

Estimating moments of Phase Geometric distribution

The mean and variance can be calculated as:

$$E(\hat{N}) = \Upsilon[I-P][P]^{-1}1^T \qquad Var(\hat{N}) = \Upsilon[I-P]^2[P]^{-2} + [I-P][P]^{-1}1^T \qquad \textbf{(8.22)}$$

Replacing equation and (8.22) with values of stationary probabilities and matrix parameter for three, four, five and six state model the mean and variance respectively are:

$$E(\hat{N}) = -0.1742 \qquad E(\hat{N}) = -8.8700e^{-5} \qquad E(\hat{N}) = -0.0074 \qquad E(\hat{N}) = -8.4700e^{-5}$$
$$Var(\hat{N}) = 0.2736 \qquad Var(\hat{N}) = 1.258e^{-5} \qquad Var(\hat{N}) = 0.0233 \qquad Var(\hat{N}) = 1.7640e^{-5}$$

8.4.5 Tabulation of frequency probabilities for PH class $(a, b, 0)$

The results obtained above are tabulated for easy comparison and interpretation. The probabilities of phase type distribution of class $(a, b, 0)$ are tabulated as:

States	Actual	PH Poisson	PH Neg Binom	PH Binomial	PH Geometric
3 state	0.0435	0.0245	$2.1865 * 10^{-19}$	$2.2457 * 10^{-102}$	$3.4694 * 10^{-18}$
4 state	0.0168	0.0184	$2.5411 * 10^{-22}$	0	$1.4476 * 10^{-24}$
5 state	0.0174	0.0147	0	$1.5888 * 10^{-102}$	0
6 state	0.0224	0.0123	$1.1294 * 10^{-22}$	0	$3.3087 * 10^{-24}$

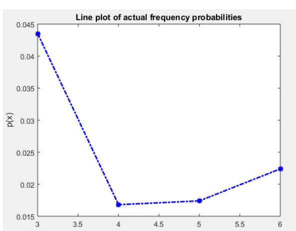
Table 8.3. Frequency probabilities of PH distributions for class $(a, b, 0)$ in different states

Table (8.3) shows that PH Poisson distribution has higher frequency probabilities than PH Negative Binomial , PH Binomial and PH Geometric distributions which are closer to the actual frequency probabilities. Table (8.4) represents p-values and Multiple R of frequency probabilities of Panjer class $(a, b, 0)$ shown in table (8.3) .

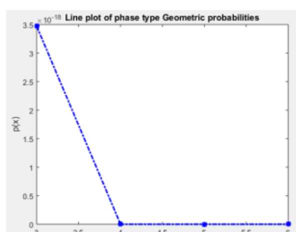
Hypothesis	P-value	Multiple R
H_0 : Estimated frequency probabilities of PH Panjer class $(a, b, 0)$ are not correlated to actual frequency probabilities		
PH-Poisson	0.0074	0.9663
PH-Neg Binom	0.1060	0.7980
PH-Binomial	0.0786	0.8347
PH-Geometric	0.1064	0.7969

Table 8.4. P-values and Multiple R for Panjer class $(a, b, 0)$ **frequency probabilities**

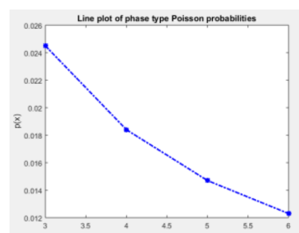
Table (8.4) indicates that the p-values for the models considered were higher than 0.05 except for PH Poisson indicating that it was the only significant model for Panjer class $(a, b, 0)$ models. It had the highest Multiple R value indicating that it provided the best fit for frequency data. The values of claim count distributions for Panjer class $(a, b, 0)$ are represented in figure (8.2) as:



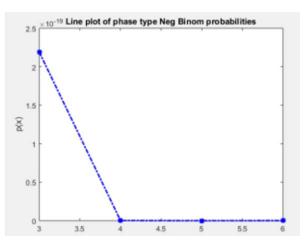
(a) Actual frequency probabilities



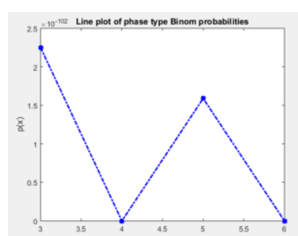
(b)PH Geometric probabilities



(c) PH Poisson Probabilities



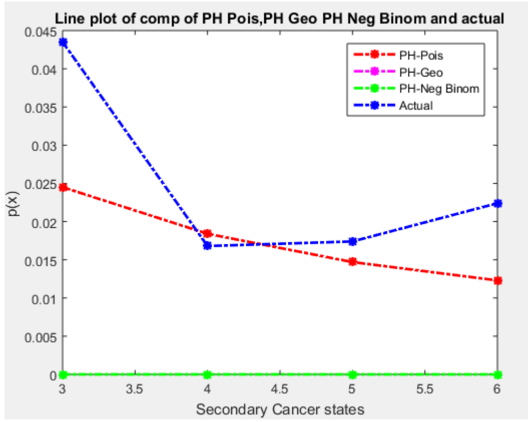
(d) PH Neg Binomial Probabilities



(e) PH Binomial Probabilities

Figure 8.2. PH Poisson, PH Geometric, PH Neg Binom, PH Binomial and Actual probabilities

From figure (8.2) it is evident that PH Poisson probabilities are similar to the actual frequency probabilities hence they are compared to determine if it provides a good fit for the data. Frequency probabilities using PH Poisson, PH Negative Binomial, PH Geometric and actual data can be compared graphically as:



(a) Comparison of class $(a, b, 0)$ probabilities

Figure 8.3. Comparison of PH Poisson and Actual probabilities

Figure (8.3) shows that PH Poisson distribution provided beNer fit for class $(a, b, 0)$ distributions as it was close to the actual frequency probabilities. This is supported by its p-values and Multiple R values in Table (8.4) as it has a p-value less than 0.05 and highest Multiple R value. The moments of different state of phase type Poisson, Negative Binomial, Geometric and Binomial distribution are tabulated as:

States	3 state	4 sate	5 state	6 state
E(N)[Pois]	1.0001	1	1.0001	1
Var (N) [Pois]	1.0001	1	1.0001	1
E(N) [NB]	6.8116	$5.0592 * e^{-04}$	0.0106	$1.5577e^{-04}$
Var (N) [NB]	11.0033	$8.9618e^{-04}$	0.0135	$2.6274e^{-04}$
E(N) [Binom]	471.0471	471	471.0471	471
Var (N) [Binom]	$6.5364 * 10^{-15}$	0	0	$1.5958 * e^{-18}$
E(N) [Geo]	-0.1742	$-8.8708 * 10^{-5}$	-0.0074	$-8.47 * 10^{-15}$
Var (N) [Geo]	0.2736	$1.258 * 10^{-5}$	0.0233	$1.7640 * 10^{-5}$

Table 8.5. Moments of PH Pois, Neg Binom, Binom and Geo distributions

Table (8.5) shows that the mean and variance of Poisson is equal which should be the case for Poisson distributions. PH Negative Binomial, PH Binomial and PH Geometric distribution indicated under dispersion as their variance was less than the actual data variance.

8.5 Estimation of frequency probabilities and its moments for class $(a, b, 1)$

Frequency probabilities for phase type Zero-truncated Poisson, Phase type Zero-truncated Binomial and Phase type Zero truncated Geometric distribution are estimated.

8.5.1 Probabilities of Phase type Zero-truncated Poisson distribution

Phase type Zero-truncated Poisson distribution is expressed as:

$$p_n = \Upsilon\Lambda^n(e^{\Lambda} - I)^{-1}\mathbb{1}^T \qquad n = 1, 2, 3, \ldots \tag{8.23}$$

Λ is a $M * M$ matrix which is be evaluated as $3 * 3$ matrix, $4 * 4$ matrix, $5 * 5$ matrix, and $6 * 6$ matrix in this research. The values of Λ and Υ have already been estimated hence equation (8.23) can be evaluated. The values of Λ and Υ are:

$$\Lambda = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \tag{8.24}$$

Replacing values of four, five and six state model in equation (8.23) the probabilities respectively are:

$$p_3 = 0.0388 \qquad p_4 = 0.0291 \qquad p_5 = 0.0233 \qquad p_6 = 0.0194 \tag{8.25}$$

Estimating moments of Phase type Zero Truncated Poisson distribution

The mean and variance of phase type Zero truncated Poisson can be evaluated using the formulas;

$$E(\hat{N}) = \Upsilon\Lambda[I - e^{-\Lambda}]^{-1}\mathbb{1}^T \qquad Var(\hat{N}) = \Upsilon(\Lambda + \Lambda^2)(I - e^{-\Lambda}) - \Lambda^2(I - e^{-\Lambda})^{-2}\mathbb{1}^T \tag{8.26}$$

Replacing equation (8.26) with values of stationary probabilities and matrix parameter for three, four, five and six state model the mean and variance respectively are:

$$E(\hat{N}) = 1.5821 \qquad E(\hat{N}) = 1.5820 \qquad E(\hat{N}) = 1.5821 \qquad E(\hat{N}) = 1.5820$$
$$Var(\hat{N}) = 0.6614 \qquad Var(\hat{N}) = 0.6613 \qquad Var(\hat{N}) = 0.6614 \qquad Var(\hat{N}) = 0.6613$$

8.5.2 Probabilities of Phase type Zero-truncated Binomial distribution

Phase type Zero-truncated Binomial distribution is expressed as:

$$p_n = \Upsilon\left[\binom{\alpha}{n}[I - Q]^n Q^{\alpha - n}\right][I - Q^{\alpha}]^{-1}\mathbb{1}^T \qquad n = 1, 2, 3, \ldots \tag{8.27}$$

P is a $(M * M)$ matrix which is evaluated as $(3 * 3)$ matrix,$(4 * 4)$ matrix,$(5 * 5)$ matrix, and $(6 * 6)$ matrix in this research. The values of P and Υ have already been estimated hence equation (8.27) can be evaluated. The values of P and Υ are:

$$P = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \tag{8.28}$$

The value of α can be estimated as from the expression of the mean for all different state models as :

$$E(\hat{N}) = \Upsilon \alpha P[I - Q^{\alpha}]^{-1} 1^{T} \qquad \alpha = 471 \tag{8.29}$$

Replacing values of three, four, five and six state model in equation (8.27) the aggregate loss probabilities respectively are:

$$p_3 = 2.2457 * e^{-102} \qquad p_4 = 0 \qquad p_5 = 1.2710 * e^{-102} \qquad p_6 = 0 \tag{8.30}$$

Estimating moments of Zero-truncated Binomial distribution

The mean and variance of phase type Zero truncated Binomial can be evaluated using the formulas;

$$E(\hat{N}) = \Upsilon \alpha P[I - Q^{\alpha}]^{-1} 1^{T} \qquad Var(\hat{N}) = \Upsilon \left[\alpha P\{Q - (Q + \alpha P)Q^{\alpha}\}\right]\left[I - Q^{\alpha}\right]^{-2} 1^{T} \tag{8.31}$$

Replacing equation (8.31) with values of stationary probabilities and matrix parameter for three state model, four state model, five state model and six state model the mean and variance respectively are:

$$E(\hat{N}) = 471.0471 \qquad E(\hat{N}) = 471 \qquad E(\hat{N}) = 471.0471 \qquad E(\hat{N}) = 471$$

$$Var(\hat{N}) = 7.1054 * e^{-15} \qquad Var(\hat{N}) = 8.6736 * e^{-19} \qquad Var(\hat{N}) = 0 \qquad Var(\hat{N}) = 8.6736 * e^{-19}$$

8.5.3 Probabilities of Phase type Zero-truncated Geometric distribution

Phase type Zero-truncated Geometric distribution is expressed as:

$$p_n = \Upsilon P(I - P)^{n-1} 1^{T} \qquad n = 1, 2, 3, \ldots \tag{8.32}$$

P is a $M * M$ matrix which is evaluated as $3 * 3$ matrix, $4 * 4$ matrix, $5 * 5$ matrix, and $6 * 6$ matrix in this research. The values of P and Υ have already been estimated hence equation (8.32) can be evaluated. The values of P and Υ are:

$$P = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \tag{8.33}$$

Replacing values of three, four, five and six state model in equation (8.32) the probabilities respectively are:

$$p_3 = 9.2519 * e^{-19} \qquad p_4 = 0 \qquad p_5 = 0 \qquad p_6 = 1.1294 * e^{-22} \tag{8.34}$$

Estimating moments of Phase type Zero Truncated Geometric distribution

The mean and variance of phase type Zero truncated Geometric can be evaluated using the formulas;

$$E(\hat{N}) = \Upsilon P^{-1} 1^T \qquad Var(\hat{N}) = \Upsilon [I - P] P^{-2} 1^T \tag{8.35}$$

Replacing equation (8.35) with values of stationary probabilities and matrix parameter for three state model, four state model, five state model and six state model the mean and variance respectively are:

$$E(\hat{N}) = 1.1262 \qquad E(\hat{N}) = 1.0000 \qquad E(\hat{N}) = 1.0127 \qquad E(\hat{N}) = 1.0000$$

$$Var(\hat{N}) = 1.38 * e^{-17} \quad Var(\hat{N}) = 1.18 * e^{-20} \quad Var(\hat{N}) = 4.33 * e^{-19} \quad Var(\hat{N}) = 6.77 * e^{-21}$$

8.5.4 Tabulation of frequency probabilities for class $(a, b, 1)$

The results obtained above can be tabulated for easy comparison and interpretation. The probabilities of phase type distribution of class $(a, b, 1)$ are tabulated as:

States	Actual	PH ZT Poisson	PH ZT Binomial	PH ZT Geometric
3 state	0.0435	0.0388	$2.2457 * e^{-102}$	$9.2519 * e^{-19}$
4 state	0.0168	0.0291	0	0
5 state	0.0174	0.0233	$1.2710 * e^{-102}$	0
6 state	0.0224	0.0194	0	$1.1294 * e^{-22}$

Table 8.6. Frequency probabilities of PH distributions for class $(a, b, 1)$ **in different states**

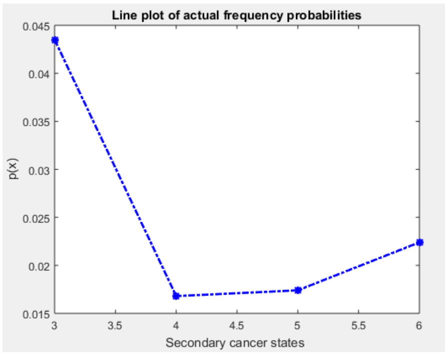
Table (8.6) shows that PH Zero-truncated Poisson distribution has higher frequency probabilities than PH Zero-truncated Binomial and PH Zero-truncated Geometric distributions and it is closer to the actual frequency probabilities. Table (8.7) represents p-values and Multiple R of frequency probabilities of PH Panjer class $(a, b, 1)$ shown in table (8.6) .

Hypothesis	P-value	Multiple R
H_0 : Estimated frequency probabilities of PH Panjer class $(a, b, 1)$ are not correlated to actual frequency probabilities		
PH-ZT Poisson	0.0073	0.9663
PH-ZT Binomial	0.0677	0.8506
PH-ZT Geometric	0.1064	0.7970

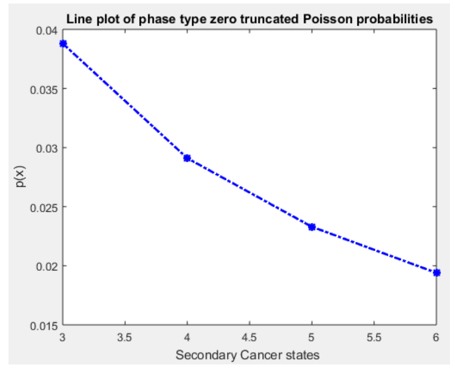
Table 8.7. P-values and Multiple R for Panjer class $(a, b, 1)$ **frequency probabilities**

Table (8.7) illustrates that the p-values for the models considered were higher than 0.05 except for PH

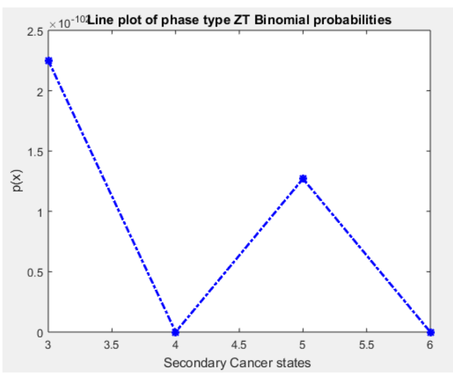
ZT Poisson indicating that it was the only model that provided a good fit for frequency probabilities for Panjer class $(a, b, 1)$. The values of phase type claim count distributions for class $(a, b, 1)$ are represented in figure (8.4) as:



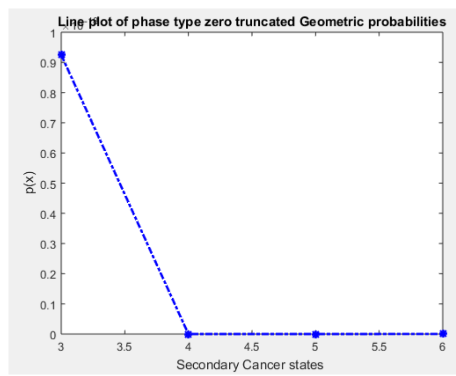
(a) Actual frequency probabilities



(b) PH ZT Poisson Probabilities



(c) PH ZT Binomial Probabilities



(d) PH ZT Geometric Probabilities

Figure 8.4. PH ZT Poisson, PH ZT Binom, PH ZT Geometric and Actual probabilities

From figure (8.4) it is evident that PH Zero-truncated Poisson probabilities are similar to the actual frequency probabilities hence they are compared to actual frequency probabilities to determine if it provides a good fit for the data. Frequency probabilities using PH ZT Poisson and actual probabilities can be represented graphically as:

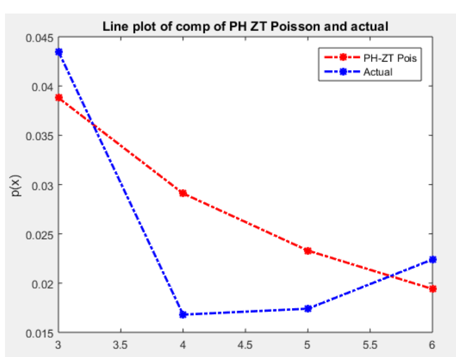


Figure 8.5. Comparison of PH ZT Poisson and Actual probabilities

Figure (8.5) shows that PH ZT Poisson distribution provided beNer fit for class $(a, b, 1)$ distributions as it was closer to the actual frequency probabilities. The moments of different state of PH zero-truncated Poisson, PH zero-truncated Binomial and PH zero-truncated Geometric distributions are tabulated as:

States	3 state	4 state	5 state	6 state
E(N)[PH ZT Pois]	1.5821	1.5820	1.5721	1.5820
Var (N) [PH ZT Pois]	0.6614	0.6613	0.6614	0.6613
E(N) [PH ZT Binom]	471.0471	471	471.0471	471
Var (N) [PH ZT Binom]	$7.1054 * e^{-15}$	$8.6736 * e^{-19}$	0	$8.6736 * e^{-19}$
E(N) [PH ZT Geo]	1.1262	1	1.0127	1
Var (N) [PH ZT Geo]	$1.3878 * e^{-27}$	$1.1858 * e^{-20}$	$4.3668 * e^{-19}$	$6.7763 * e^{-21}$

Table 8.8. Moments of PH ZT Pois, ZT Binom and PH ZT Geo distributions

Table (8.8) shows that modeling the data with PH Zero-truncated Poisson distribution, PH Zero-truncated Binomial distribution and PH Zero-truncated Geometric distribution indicates under dispersion as the variance is less than variance of the actual data.

8.6 Estimation of frequency probabilities and moments of phase type Poisson Lindley distributions

Frequency probabilities and moments of phase type one parameter Poisson Lindley distribution and two parameter Poisson Lindley distribution are estimated.

8.6.1 Probabilities of Phase type one parameter Poisson Lindley distribution

Phase type one parameter Poisson distribution is expressed as:

$$p_n = \Upsilon\Gamma^2(I+\Gamma)^{-(n+3)}\{(n+2)I+\Gamma\}1^T \qquad n = 0, 1, 2, \ldots \tag{8.36}$$

The values of Γ and Υ have already been estimated hence equation (8.36) can be evaluated. The values of Γ and Υ are:

$$\Gamma = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \tag{8.37}$$

Replacing values of three, four, five and six state model in equation (8.36) the probabilities respectively are:

$$p_3 = 0.0169 \qquad p_4 = 0.0125 \qquad p_5 = 0.01 \qquad p_6 = 0.0083 \tag{8.38}$$

Estimating moments of Phase one parameter Poisson Lindley distribution

The mean and variance can be calculated as:

$$E(\hat{N}) = \Upsilon(2I+\Gamma)\left[\Gamma(\Gamma+I)\right]^{-1}\mathbf{1}^T \qquad Var(\hat{N}) = \Upsilon(2I+4\Gamma+\Gamma^2)(\Gamma+I)^{-2}+(2I+\Gamma)[\Gamma(\Gamma+I)]^{-1}\mathbf{1}^T \quad (8.39)$$

Replacing equation (8.39) with values of stationary probabilities and matrix parameter for three state model, four state model, five state model and six state model the mean and variance respectively are:

$$E(\hat{N}) = 1.8622 \quad E(\hat{N}) = 1.5017 \quad E(\hat{N}) = 1.5099 \quad E(\hat{N}) = 1.5$$
$$Var(\hat{N}) = 4.3527 \quad Var(\hat{N}) = 3.2554 \quad Var(\hat{N}) = 3.2727 \quad Var(\hat{N}) = 3.2502$$

8.6.2 Probabilities of Phase type two parameter Poisson Lindley distribution

Phase type two parameter Poisson distribution is expressed as:

$$p_n = \Upsilon\Gamma^2(I+\Gamma)^{-(n+2)}\left[I+[(\alpha+n)I][\alpha\Gamma+I]^{-1}\right]\mathbf{1}^T \qquad n = 0,1,2,\ldots \quad (8.40)$$

where $\alpha > 0$,, Γ is a $(M * M)$ matrix.
The values of Γ and Υ have already been estimated hence equation (8.40) can be evaluated. The values of Γ and Υ are:

$$\Gamma = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \quad (8.41)$$

The value of α can be estimated from the expression of the mean as:

$$E(N) = \Upsilon(2I+\Gamma\alpha)[\Gamma(\Gamma\alpha+I)]^{-1}\mathbf{1}^T \quad \alpha_3 = -0.0667 \quad \alpha_4 = -0.0499 \quad \alpha_5 = -0.0399 \quad \alpha_6 = -0.0333 \quad (8.42)$$

Replacing values of four, five and six state model in equation (8.40) the probabilities respectively are:

$$p_3 = 0.0187 \qquad p_4 = 0.0125 \qquad p_5 = 0.01 \qquad p_6 = 0.0083 \qquad (8.43)$$

Estimating moments of Phase type two parameter Poisson Lindley distribution

The mean and variance can be calculated as:

$$E(\hat{N}) = \Upsilon(2I+\Gamma\alpha)[\Gamma(\Gamma\alpha+I)]^{-1}\mathbf{1}^T \quad (8.44)$$

$$Var(\hat{N}) = \Upsilon[\alpha\Gamma+2I][\Gamma(\alpha\Gamma+I)]^{-1}+[2(\alpha\Gamma+3I)][\Gamma^2(\alpha\Gamma+I)]^{-1}-\left[(2I+\Gamma\alpha)][\Gamma(\Gamma\alpha+I)]^{-1}\right]^2\mathbf{1}^T \quad (8.45)$$

Replacing equation (8.45) with values of stationary probabilities and matrix parameter for three state model, four state model, five state model and six state model the mean and variance respectively are:

$$E(\hat{N}) = 2.0717 \qquad E(\hat{N}) = 2.0525 \qquad E(\hat{N}) = 2.0416 \qquad E(\hat{N}) = 2.0344$$
$$Var(\hat{N}) = 4.0668 \qquad Var(\hat{N}) = 4.0498 \qquad Var(\hat{N}) = 4.0398 \qquad Var(\hat{N}) = 4.0333$$

8.6.3 Tabulation of frequency probabilities for PH Poisson Lindley distributions

The results obtained above are tabulated for easy comparison and interpretation. The probabilities of phase type Poisson Lindley distributions are tabulated as:

Distributions	3 state	4 state	5 state	6 state
PH-OPPL	0.0169	0.0125	0.01	0.0083
PH-TPPL	0.0187	0.0125	0.01	0.0083
Actual	0.0435	0.0168	0.0174	0.0224

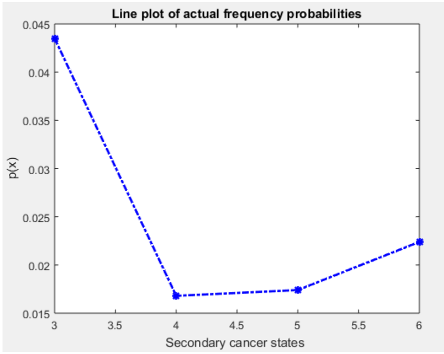
Table 8.9. Frequency probabilities of PH Poisson Lindley distributions in different states

Table (8.9) shows that PH one parameter Poisson Lindley and PH two parameter Poisson Lindley distributions had higher frequency probabilities which are closer to actual frequency distributions hence they will be compared to determine which between them provides the best fit. Table (8.10) represents p-values and Multiple R of frequency probabilities of phase type OPPL and phase type TPPL distributions shown in table (8.9) .

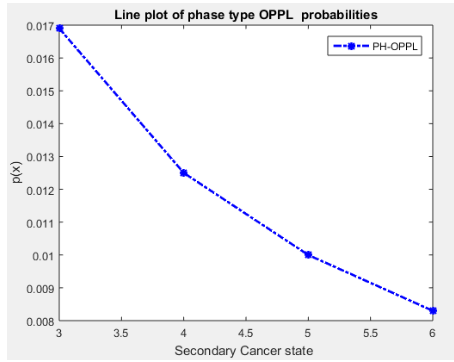
Hypothesis	P-value	Multiple R
H_0 : Estimated frequency probabilities of PH Poisson Lindley distributions are not correlated to actual frequency probabilities		
PH-OPPL	0.0070	0.9675
PH-TPPL	0.0045	0.9757

Table 8.10. P-values and Multiple R for PH Poisson Lindley frequency probabilities

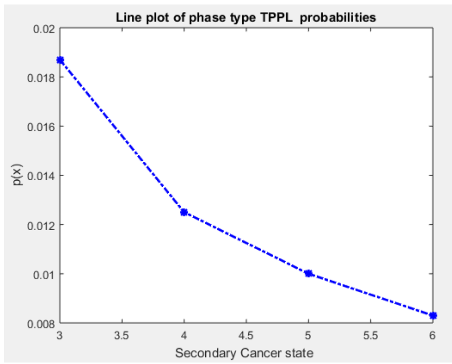
Table (8.10) shows that the p-values for all models were lesser than 0.05 indicating that they provided a good fit for frequency probabilities however PH-TPPL had the least p-value and highest Multiple R indication that it provided the best fit for frequency data. The values of frequency distributions for phase type Poisson Lindley distributions are represented graphically in figure (8.6) as:



(a) Actual frequency probabilities



(b) PH OPPL Probabilities



(c) PH TPPL Probabilities

Figure 8.6. PH-OPPL, PH-TPPL and Actual probabilities

Figure (8.6) indicates that phase type one parameter Poisson Lindley and phase type two parameter Poisson Lindley are similar to actual frequency data hence they will be compared to determine which provides a beNer fit. Frequency probabilities using , PH-OPPL, PH-TPPL and actual data can be represented graphically as:

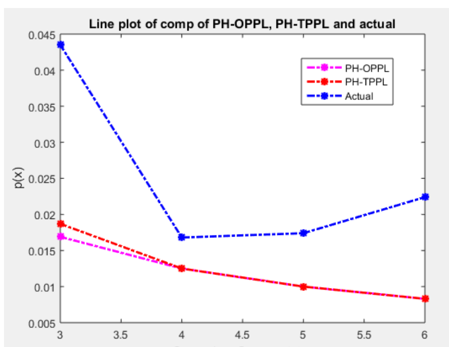


Figure 8.7. Comparison of PH Poisson Lindley distributions and Actual probabilities

Figure (8.7) shows that PH two parameter Poisson Lindley distribution provided beNer fit as it is closer to the actual frequency probabilities and has the lowest p-value. The moments of different states of phase type one parameter Poisson Lindley and phase type two parameter Poisson Lindley distributions are tabulated as:

States	3 state	4 state	5 state	6 state
E(N) [PH-OPPL]	1.8622	1.5017	1.5099	1.5
Var (N) [PH-OPPL]	4.3527	3.2554	3.2727	3.2502
E(N) [PH-TPPL]	2.0717	2.0525	2.0416	2.0344
Var (N) [PH-TPPL]	4.0668	4.0498	4.0398	4.0333

Table 8.11. Moments of PH-OPPL and PH-TPPL distribution

Table (8.11) shows that modeling the data with PH one parameter Poisson Lindley distribution and PH two parameter Poisson Lindley indicates under dispersion as its variance is less than variance of the actual data.

8.7 Estimation of frequency probabilities and moments of phase type Zero-truncated Poisson Lindley distributions

Frequency probabilities and moments of phase type Zero-truncated one parameter Poisson Lindley distribution and Zero-truncated two parameter Poisson Lindley distribution are estimated.

8.7.1 Probabilities of Phase type Zero-truncated one parameter Poisson Lindley distribution

Phase type Zero-truncated one parameter Poisson distribution is expressed as:

$$p_n = \Upsilon\Lambda^2[\Lambda^2 + 3\Lambda + I]^{-1}\left[(n+2)I + \Lambda\right][(\Lambda + I)^n]^{-1}\mathbf{1}^T \qquad n = 1, 2, \ldots \qquad \textbf{(8.46)}$$

The values of Λ and Υ have already been estimated hence equation (8.46) can be evaluated. The values of Λ and Υ are:

$$\Lambda = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \qquad \textbf{(8.47)}$$

Replacing values of four, five and six state model in equation (8.46) the probabilities respectively are:

$$p_3 = 0.0267 \qquad p_4 = 0.02 \qquad p_5 = 0.016 \qquad p_6 = 0.0133 \qquad \textbf{(8.48)}$$

Estimating moments of Phase Zero-truncated one parameter Poisson Lindley distribution

The mean and variance can be calculated as:

$$E(\hat{N}) = \Upsilon\left[(\Lambda+I)^2(2I+\Lambda)\right]\left[\Lambda(\Lambda^2+3\Lambda+I)\right]^{-1}\mathbb{1}^T \quad \textbf{(8.49)}$$

$$Var(\hat{N}) = \Upsilon\left[(\Lambda+I)^2(\Lambda^2+4\Lambda+6I)\right]\left[\Lambda^2(\Lambda^2+3\Lambda+I)\right]^{-1} - \left[\Upsilon\left[(\Lambda+I)^2(2I+\Lambda)\right]\left[\Lambda(\Lambda^2+3\Lambda+I)\right]^{-1}\mathbb{1}^T\right]^2 \mathbb{1}^T \quad \textbf{(8.50)}$$

Replacing equation (8.50) with values of stationary probabilities and matrix parameter for three state model, four state model, five state model and six state model the mean and variance respectively are:

$$E(\hat{N}) = 2.3043 \qquad E(\hat{N}) = 2.3996 \qquad E(\hat{N}) = 2.3967 \qquad E(\hat{N}) = 2.4000$$
$$Var(\hat{N}) = 4.3177 \qquad Var(\hat{N}) = 3.0443 \qquad Var(\hat{N}) = 3.0874 \qquad Var(\hat{N}) = 3.0402$$

8.7.2 Probabilities of Phase type Zero-truncated two parameter Poisson Lindley distribution

Phase type Zero-truncated two parameter Poisson distribution is illustrated as:

$$p_n = \Upsilon\left[\Lambda^2[\Lambda^2\alpha+\Lambda\alpha+2\Lambda+I]^{-1}\left[(n+1)I+\alpha(\Lambda+I)\right](\Lambda+I)^{-n}\right]\mathbb{1}^T \qquad n = 1, 2, \ldots \quad \textbf{(8.51)}$$

where $\alpha > 0$, Λ is a $(M * M)$ matrix.
The values of Λ and Υ have already been estimated hence equation (8.51) can be evaluated. The values of Λ and Υ are:

$$\Lambda = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \Upsilon = \begin{bmatrix} 0.2122 & 0.0535 & 0.7343 \end{bmatrix} \quad \textbf{(8.52)}$$

The value of α can be estimated from the expression of the mean as:

$$E(N) = \Upsilon\left[(\Lambda+I)^2[\Lambda\alpha+2I]\right]\left[\Lambda[\Lambda^2\alpha+\Lambda\alpha+2\Lambda+I]\right]^{-1}\mathbb{1}^T$$
$$\alpha_3 = -0.0999 \qquad \alpha_4 = -0.0749 \qquad \alpha_5 = -0.0599 \qquad \alpha_6 = -0.0499 \quad \textbf{(8.53)}$$

Replacing values of four, five and six state model in equation (8.51) the probabilities respectively are:

$$p_3 = 0.0197 \qquad p_4 = 0.0162 \qquad p_5 = 0.0130 \qquad p_6 = 0.0109 \quad \textbf{(8.54)}$$

Estimating moments of Phase type ZT two parameter Poisson Lindley distribution

The mean and variance can be calculated as:

$$E(\hat{N}) = \Upsilon\left[(\Lambda+I)^2(\Lambda\alpha+2I)\right]\left[\Lambda(\Lambda^2\alpha+\Lambda\alpha+2\Lambda+I)\right]^{-1}\mathbb{1}^T \quad \textbf{(8.55)}$$

$$Var(\hat{N}) = \vec{\Upsilon}\left[(\Lambda+I)^2(\Lambda^3\alpha^2+(\alpha+5)\Lambda^2\alpha+(4\alpha+6)\Lambda+2I)\right]\left[\Lambda^2(\Lambda^2\alpha+\Lambda\alpha 2+\Lambda+I)^2\right]^{-1}\vec{1}^T \quad \textbf{(8.56)}$$

Replacing equation (8.55) and (8.56) with values of stationary probabilities and matrix parameter for three state model, four state model, five state model and six state model the mean and variance respectively are:

$E(\hat{N}) = 2.5653$ $\quad E(\hat{N}) = 2.7010$ $\quad E(\hat{N}) = 2.6894$ $\quad E(\hat{N}) = 2.6896$

$Var(\hat{N}) = 11.1226$ $\quad Var(\hat{N}) = 10.2847$ $\quad Var(\hat{N}) = 10.3880$ $\quad Var(\hat{N}) = 10.4145$

8.7.3 Tabulation of frequency probabilities for PH Zero-truncated Poisson Lindley distributions

The results obtained above are tabulated for easy comparison and interpretation. The probabilities of phase type Zero-truncated Poisson Lindley distributions are tabulated as:

Distributions	3 state	4 state	5 state	6 state
PH-ZTOPPL	0.0267	0.0200	0.0160	0.0133
PH-ZTTPPL	0.0197	0.0162	0.0130	0.0109
Actual	0.0435	0.0168	0.0174	0.0224

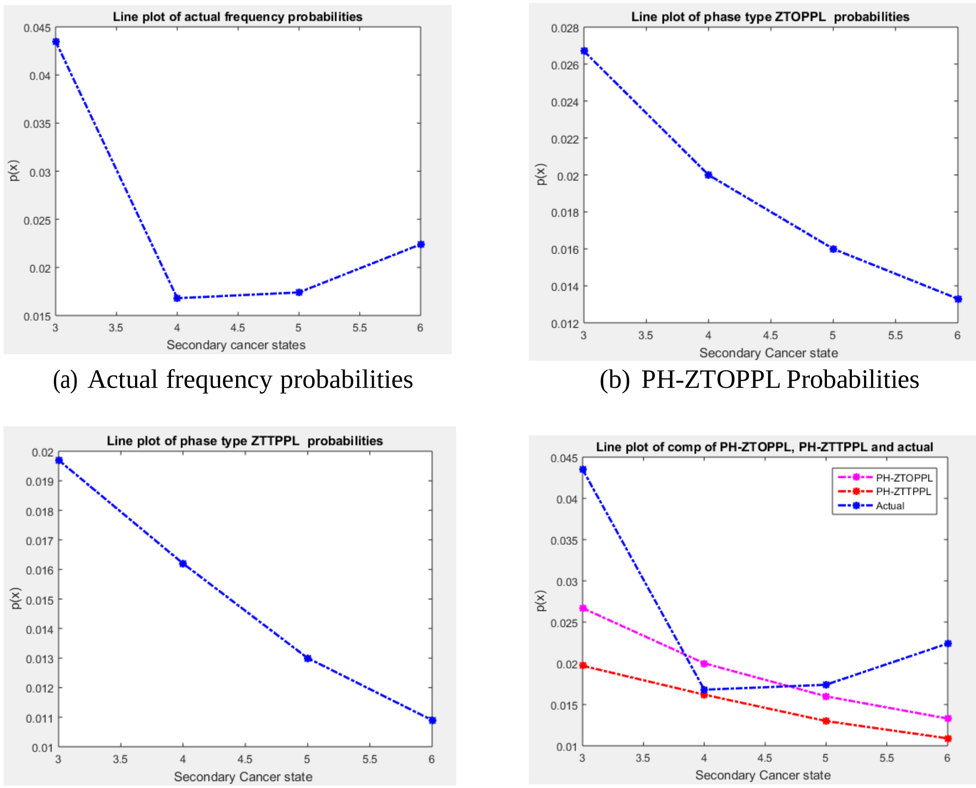
Table 8.12. Frequency probabilities of PH ZT Poisson Lindley distributions in different states

Table (8.12) shows that PH Zero-truncated one parameter Poisson Lindley and PH Zero-truncated two parameter Poisson Lindley distributions has higher frequency probabilities which are closer to actual frequency distributions hence they are compared to determine which between them provides the best fit. Table (8.13) represents p-values and Multiple R of frequency probabilities of phase type ZT-OPPL and phase type ZT-TPPL distributions shown in table (8.12) .

Hypothesis	P-value	Multiple R
H_0 : Estimated claim count probabilities of PH ZT Poisson Lindley distributions are not correlated to actual frequency probabilities		
PH-ZT OPPL	0.0074	0.9663
PH-ZT TPPL	0.0108	0.9567

Table 8.13. P-values and Multiple R for PH ZT Poisson Lindley frequency probabilities

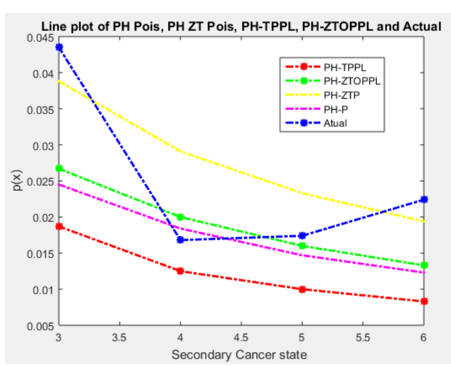
Table (8.13) shows that the p-values for all models were lesser than 0.05 indicating that they provided a good fit for frequency probabilities however PH-ZT OPPL had the least p-value and highest Multiple R indicating that it provided the best fit for frequency data. The values of frequency distrib utions for Zero-truncated Poisson Lindley distribution are represented graphically in figure (8.8) as:



(a) Actual frequency probabilities
(b) PH-ZTOPPL Probabilities
(c) PH-ZTTPPL Probabilities
(d) Comp PH-ZTOPPL and PH-ZTTPPL Probabilities

Figure 8.8. PH-ZTOPPL, PH-ZTTPPL and Actual probabilities

From figure (8.8) it is evident that PH Zero-truncated one parameter Poisson Lindley distributions and PH Zero-truncated two parameter Poisson Lindley distribution probabilities are similar to the actual frequency probabilities hence it they compared to actual frequency probabilities determine which distribution provides the best fit for the data. Frequency probabilities using PH Zero-truncated Poisson Lindley distributions and actual data can be represented graphically as:



(a) Comparison frequency Probabilities

Figure 8.9. Comparison of PH Zero-truncated Poisson Lindley and Actual probabilities

Figure (8.9) shows that PH Zero-truncated two parameter Poisson distribution provided beNer fit as it was closer to the actual frequency probabilities and also had the lowest p-value and Multiple R.

The moments of different state of phase type Zero-truncated one parameter Poisson Lindley and phase type two parameter Poisson Lindley distributions are tabulated as:

States	3 state	4 state	5 state	6 state
E(N) [PH-ZTOPPL]	2.3043	2.3996	2.3967	2.4000
Var (N) [PH-ZTOPPL]	4.3177	3.0443	3.0874	3.0402
E(N) [PH-ZTTPPL]	2.5653	2.7010	2.6894	2.6896
Var (N) [PH-ZTTPPL]	11.1226	10.2847	10.3880	10.4145

Table 8.14. Moments of PH-ZTOPPL and PH-ZTTPPL distribution

Table (8.14) shows that modeling the data with PH Zero-truncated one parameter Poisson Lindley and PH Zero-truncated two parameter Poisson Lindley distributions exhibits under dispersion as its variance is less than variance of the actual data.

8.8 Estimation of severity distribution

The parameters of both discrete and continuous severity distributions are estimated. Continuous distribution taken into account are Weibull distribution, Generalized Pareto, and Pareto distribution. The continuous distributions are arithmetized using method of mass rounding and the span chosen is $h = 250000$, $h = 200000$, $h = 300000$, $h = 1200000$ for three state model, four state model, five state model and six state models respectively. Discrete distribution considered in this research are one-parameter Poisson lindley distribution and two-parameter Poisson Lindley distribution. In this section parameters of the distributions are estimated as well as the severity probabilities.

8.8.1 Estimation of parameters and discretization of Weibull distribution

From chapter seven estimate of α is expressed as:

$$\alpha = \frac{E(x)}{\Gamma(1+\frac{1}{\beta})} \qquad (8.57)$$

The estimate of α requires β to be estimated first as:

$$Var(x) = \frac{[E(x)]^2}{[\Gamma(1+\frac{1}{\beta})]^2} * \Gamma(1+\frac{2}{\beta}) - [E(x)]^2 \qquad \frac{Var(x) - [E(x)]^2}{[E(x)]^2} = \frac{\Gamma(1+\frac{2}{\beta})}{[\Gamma(1+\frac{1}{\beta})]^2} \qquad (8.58)$$

The values of β can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.58) as:

3 state = 0.47 4 state = 0.815 5 state = 9.1 6 state = 0.842

Hence the value of α can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.57) as:

3 state = 150507.7436   4 state = 240521.1121   5 state = 432431.6298   6 state = 1539559.443

Weibull distribution is discretized using the following formula:

$$f_s = F_j\left(xh+\frac{h}{2}\right) - F_j\left(xh-\frac{h}{2}\right)$$

The resulting probabilities estimates for three, four, five and six state models for Weibull distribution are:

$$f_3 = 0.0171 \quad f_4 = 0.0216 \quad f_5 = 0 \quad f_6 = 0.0231$$

8.8.2 Estimation of parameters and discretization of Generalized Pareto distribution

From chapter six Generalized Pareto distribution can also be expressed as;

$$f(x) = \frac{1}{e}\left[1-\frac{d(x-f)}{e}\right]^{\frac{1}{d}-1} \qquad X = e[1-exp(-dY)]/d.$$

The moment estimators of Generalized Pareto distribution are:

$$E(x) = f + \frac{e}{1+d} \quad Var(x) = \frac{e^2}{(1+d)^2(1+2d)} \quad Skew(x) = \frac{2(1-d)(1+2d)^{0.5}}{1+3d} \tag{8.59}$$

The value of d can be calculated from the skewness formula, hence e estimated as follows:

$$e = S.D(1+d)(1+2d)^{0.5} \tag{8.60}$$

The value of f can be calculated as:

$$f = \bar{x} - \frac{e}{1+d} \tag{8.61}$$

The values of d can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.59) as:

3 state = 17.59   4 state = 5.079   5 state = 2.13   6 state = 0.268

The values of e can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.60) as:

3 state = 107440000   4 state = 10269000   5 state = 4179000   6 state = 4901100

The values of f can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.61) as:

3 state = 339721.7777   4 state = 268984   5 state = 409701.2222   6 state = 1684997.555

The cumulative distribution function is expressed as:

$$F_X = 1 - \left[1-\frac{d(x-f)}{e}\right]^{\frac{1}{d}} \tag{8.62}$$

The resulting probabilities estimates for three, four, five and six state models for Generalized Pareto distribution are:

$$f_3 = 0.0027 \quad f_4 = 0.0280 \quad f_5 = 0.1108 \quad f_6 = 0.1176$$

8.8.3 Estimation of parameters and discretization of Pareto distribution

From chapter six Pareto distribution can also be expressed as;

$$f(x) = \frac{\alpha\beta^{\alpha}}{(x+\beta)^{\alpha+1}} \tag{8.63}$$

The value of β is estimated as follows:

$$E(x) = \frac{\beta}{\alpha - 1} \quad \alpha > 1 \qquad E(x^2) = \frac{2\beta^2}{(\alpha-1)(\alpha-2)}$$

We can estimate β by using these two expressions.

$$\frac{E(x^2)}{[E(x)]^2} = \frac{\frac{2\beta^2}{(\alpha-1)(\alpha-2)}}{\frac{\beta^2}{(\alpha-1)^2}} = \frac{2\beta^2}{(\alpha-1)(\alpha-2)}\frac{(\alpha-1)^2}{\beta^2} = \frac{2(\alpha-1)}{(\alpha-2)} \tag{8.64}$$

The values of α can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.64) as:

3 state = 2.285714286 4 state = 2.789019363 5 state = 3.962650275 6 state = 2.824467301

Hence the value of β can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.63) as:

3 state = 436786.4286 4 state = 481219.3733 5 state = 1213804.401 6 state = 3074224.767

The cumulative distribution function of Pareto distribution is expressed as:

$$F(x) = 1 - \left(\frac{\beta}{\beta + x}\right)^{\alpha} \tag{8.65}$$

The resulting probabilities estimates for three, four, five and six state models for Pareto distribution are:

$$f_3 = 0.0156 \quad f_4 = 0.0166 \quad f_5 = 0.0184 \quad f_6 = 0.0178$$

8.8.4 Estimation of parameters and probability of one-parameter Poisson Linldey distribution

The probability mass function of one parameter Poison Lindley is expressed as:

$$f(x) = \frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}} \qquad x = 0, 1, 2, \ldots \tag{8.66}$$

The value of x in this research considering discrete distributions if sum of the severity values for each particular multi-state model. The value of θ is calculated as:

$$\hat{\theta} = \frac{-(\bar{X}-1) + \sqrt{(\bar{X}-1)^2 + 8\bar{X}}}{2\bar{X}} \tag{8.67}$$

The values of θ can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.67) as:

3 state = $5.8871 * e^{-6}$ 4 state = $7.4353 * e^{-6}$ 5 state = $4.8816 * e^{-6}$ 6 state = $1.1869 * e^{-6}$
For three, four, five and six state models the probabilities are evaluated respectively as:

$$f(x) = \frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}} \qquad f(x)_3 = 1.6142 * e^{-12} \qquad f(x)_4 = 2.0387 * e^{-12}$$

$$f(x)_5 = 1.3383 * e^{-12} \qquad f(x)_6 = 3.2559 * e^{-13}$$

8.8.5 Estimation of parameters and probability of two-parameter Poisson Linldey distribution

The probability mass function of one parameter Poison Lindley is expressed as:

$$f(x) = \frac{\theta^2}{(\theta+1)^{x+2}}\left[1+\frac{\alpha+x}{\alpha\theta+1}\right] \qquad x = 0,1,2,....;\theta > 0; \alpha\theta > -1 \tag{8.68}$$

Expectation of two-parameter Poisson Lindley is expressed as:

$$\bar{X} = \frac{\alpha\theta+2}{\theta(\alpha\theta+1)} \tag{8.69}$$

Let $\frac{\mu_2' - \mu_1'}{[\mu_1']^2}$ be k , it can be expressed as:

$$k = \frac{2(b+3)(b+1)}{(b+2)^2} \qquad 0 = (2-k)b^2 + (8-4k)b + (6-4k) \tag{8.70}$$

The value of k can be calculated hence equation (8.70) can be solved. The value of α and θ can be expressed as:

$$\bar{X} = \frac{b+2}{\theta(b+1)} \qquad \theta = \frac{b+2}{(b+1)\bar{X}} \qquad \hat{\alpha} = \frac{b(b+1)\bar{X}}{b+2} \tag{8.71}$$

The values of b can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.71) as:

3 state = -2.5345 4 state = -2.888268 5 state = -3.400947 6 state = -2.908
The values of θ can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.71) as:

3 state = $1.02531 * e^{-6}$ 4 state = $1.74885 * e^{-6}$ 5 state = $1.4242 * e^{-6}$ 6 state = $2.82428 * e^{-7}$
The values of α can be calculated for 3 state, 4 state, 5 state and 6 state models from equation (8.71) as:

3 state = -2471929.869 4 state = -1651525.057 5 state = -2387971.016 6 state = -10296424.92
For three, four, five and six state models the probabilities are evaluated respectively as:

$$f(x) = \frac{\theta^2}{(\theta+1)^{x+2}}\left[1+\frac{\alpha+x}{\alpha\theta+1}\right] \qquad f(x)_3 = 1.7453 * e^{-8} \qquad f(x)_4 = 1.8066 * e^{-8}$$

$$f(x)_5 = 5.7517 * e^{-9} \qquad f(x)_6 = 2.8090 * e^{-9}$$

8.8.6 Tabulation of severity data results

The results obtained above can be tabulated for easy comparison and interpretation. The probabilities of continuous and discrete severity distributions are tabulated as:

States	Actual	Weibull	Gen Pareto	Pareto	One par Pois	Two par Pois
3 state	0.0046	0.0171	0.0027	0.0156	$1.6142 * e^{-12}$	$1.7453 * e^{-8}$
4 state	0.0664	0.0216	0.0280	0.0166	$2.0387 * e^{-12}$	$1.8066 * e^{-8}$
5 state	0.0794	0	0.1108	0.0184	$1.3383 * e^{-12}$	$5.7517 * e^{-9}$
6 state	0.1551	0.0231	0.1176	0.0178	$3.2559 * e^{-13}$	$2.8090 * e^{-9}$

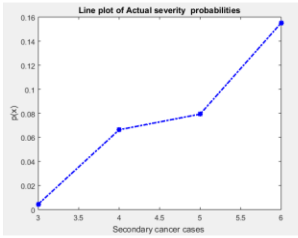
Table 8.15. Severity probabilities of continuous and discrete distributions

Table (8.15) shows Generalized Pareto had relatively higher probabilities compared to Weibull, one parameter Poisson Lindley and two parameter Poisson Lindley distributions. Consequently, Generalized Pareto had probabilities closer to the actual probabilities hence they will be compared to determine if it provides a good fit. Table (8.16) represents p-values and Multiple R of severity probabilities shown in table (8.15) .

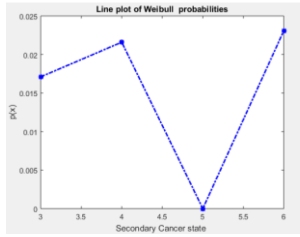
Hypothesis	P-value	Multiple R
H_0 : Estimated severity probabilities of are not correlated to actual severity probabilities		
Weibull	0.1360	0.7599
Gen Pareto	0.0153	0.9450
Pareto	0.0721	0.8442
OPPL	0.3413	0.5458
TPPL	0.4478	0.4493

Table 8.16. P-values and Multiple R for severity probabilities

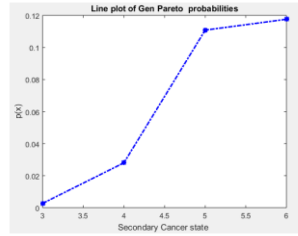
Table (8.16) indicates that the p-values for the models considered were higher than 0.05 except for Generalized Pareto distribution probabilities indicating that it provided a good fit for severity probabilities. The data of severity distribution can be represented as illustrated in figure (8.10) as:



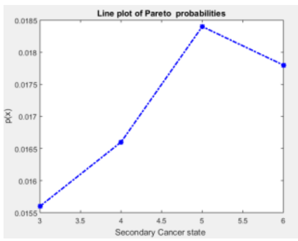
(a) Actual severity probabilities



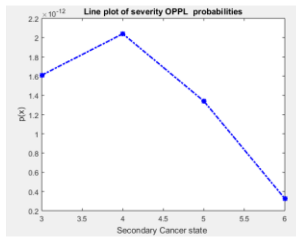
(b) Weibull Probabilities



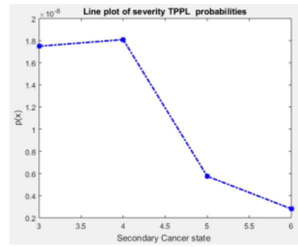
(c) Generalized Pareto Probabilities



(d) Pareto Probabilities



(e) One parameter Poisson Lindley Probabilities



(f) Two parameter Poisson Lindley probabilities

Figure 8.10. Comparison of Weibull, Generalized Pareto , Pareto, One parameter Poisson Lindley, Two parameter Poisson Lindley and Actual probabilities

Figure (8.10) shows that severity probabilities using Generalized Pareto and actual probabilities were similar hence they will be compared to determine if it provides a good fit for the severity probabilities. Generalized Pareto and actual severity probabilities can be represented graphically as:

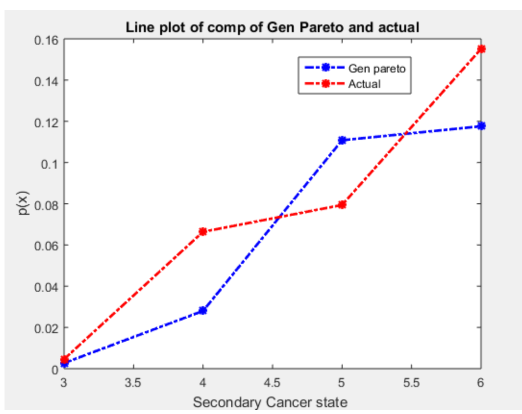


Figure 8.11. Comparison of Weibull, Generalized Pareto , Pareto and Actual probabilities

Figure (8.11) shows that Generalized Pareto provided a beNer fit for severity probabilities which is supported by its low p-values and higher R-statistic.

8.9 Estimation Aggregate loss probabilities using phase type Panjer recursive model for class $(a, b, 0)$

Aggregate loss probabilities are estimated using PH Panjer recursive model for class $(a, b, 0)$ and $(a, b, 1)$. PH Panjer recursive model for class $(a, b, 0)$ is expressed as:

$$\vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(A + B\frac{x}{j}\right)[I - Ay(0)]^{-1} \tag{8.72}$$

where $\vec{Z}(0) = \vec{Y}P_0$

8.9.1 Aggregate Loss probabilities of phase type Poisson and severity distribution

The initial condition, matrices A and B of phase type Poisson distribution are calculated as:

$$P_0 = e^{-\Lambda} \qquad A = 0 \qquad B = \Lambda$$

PH Panjer recursive model when claim frequency distribution is PH Poisson distribution is expressed as:

$$\vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\,\Lambda\frac{x}{j} \tag{8.73}$$

For three state multi-state model the initial condition is evaluated as:

$$P_0 = \begin{bmatrix} 0.4155 & -0.0651 & 0.0174 \\ 0 & 0.6745 & -0.3066 \\ 0 & 0 & 0.3679 \end{bmatrix} \qquad \vec{Y}P_0 = \begin{bmatrix} 0.0997 & 0.0251 & 0.2431 \end{bmatrix}$$

The value of A is constant and value of $B = \Lambda$ hence for three state model they can be evaluated as:

$$A = 0 \qquad B = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix}$$

Replacing values of three, four, five and six state model and severity probabilities in equation (8.72) the aggregate loss probabilities are:

States	Actual	PH Poi-Wei	PH Poi-GP	PH Poi-Par	PH Poi-OPPL	PH Poi-TPPL
3 state	0.0002	0.0065	0.0009	0.0059	$5.9232 * e^{-13}$	$6.4383 * e^{-9}$
4 state	0.0011	0.0083	0.0109	0.0063	$7.5031 * e^{-13}$	$6.6572 * e^{-9}$
5 state	0.0014	0	0.0503	0.0070	$4.9299 * e^{-13}$	$2.1154 * e^{-09}$
6 state	0.0035	0.0089	0.0540	0.0068	$1.1990 * e^{-13}$	$1.0335 * e^{-9}$

Table 8.17. Aggregate loss probabilities of phase type Poisson and severity distributions

Table (8.17) shows that actual aggregate losses have a similar trend as aggregate losses estimated using PH Poisson and Generalized Pareto.

8.9.2 Aggregate Loss probabilities of phase type Negative Binomial and severity distribution

The initial condition, matrices A and B of phase type Negative Binomial distribution are calculated as:

$$P_0 = [1 - Q]^{\alpha} \qquad A = Q \qquad B = (\alpha - 1)Q$$

Phase type panjer recursive model when claim frequency distribution is Phase type Negative Binomial distribution is expressed as:

$$\vec{Z}(j) = \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left[Q + [(\alpha-1)Q]\frac{x}{j}\right][I - Qy(0)]^{-1} \tag{8.74}$$

For three state multi-state model the initial condition is evaluated as:

$$P_0 = \begin{bmatrix} 0.0002 & 0.0002 & 0 \\ 0 & 0.1350 & 0.1350 \\ 0 & 0 & 0 \end{bmatrix} \quad \Upsilon P_0 = \begin{bmatrix} 0.0001 & 0.0082 & 0.0081 \end{bmatrix}$$

The value of $A = Q$ and value $B = (\alpha - 1)Q$ hence for three state they can be evaluated as :

$$A = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad B = \begin{bmatrix} 2.6349 & 0.3651 & 0 \\ 0 & 1.1814 & 1.8186 \\ 0 & 0 & 3.0000 \end{bmatrix}$$

Replacing values of three, four, five and six state model and severity probabilities in equation (8.74) the aggregate loss probabilities are:

States	Actual	PH NB-Wei	PH NB-GP	PH NB-Par	PH NB-OPPL	PH NB-TPPL
3 state	0.0002	0.0013	$1.8214 * e^{-4}$	0.0012	$1.0556 * e^{-13}$	$8.4721 * e^{-9}$
4 state	0.0011	$2.4268 * e^{-7}$	$3.2084 * e^{-7}$	$1.8365 * e^{-7}$	$2.1435 * e^{-17}$	$1.9019 * e^{-13}$
5 state	0.0014	0	0.0010	$1.3172 * e^{-4}$	$9.1325 * e^{-15}$	$3.9188 * e^{-11}$
6 state	0.0035	$6.3532 * e^{-7}$	$4.0625 * e^{-6}$	$4.832 * e^{-7}$	$8.4721 * e^{-18}$	$7.3027 * e^{-14}$

Table 8.18. Aggregate loss probabilities of phase type Neg Binom and severity distributions

Table (8.18) shows that the trend of actual aggregate losses are different from aggregate losses estimated using PH Negative Binomial.

8.9.3 Aggregate Loss probabilities of phase type Geometric and severity distribution

The initial condition, matrices A and B of phase type Geometric distribution are calculated as:

$$P_0 = P \qquad A = [I - P] \qquad B = 0$$

Phase type panjer recursive model when claim frequency distribution is Phase type Geometric distribution is expressed as:

$$Ƶ(j) = \sum_{x=1}^{j} y(x)Ƶ(j-x)[I-P][I-[I-P]y(0)]^{-1} \qquad \textbf{(8.75)}$$

For three state multi-state model the initial condition is evaluated as:

$$P_0 = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix} \quad ϒP_0 = \begin{bmatrix} 0.2108 & 0.0530 & 0.7364 \end{bmatrix}$$

The value of $A = [I - P]$ and value $B = 0$ hence for three state they can be evaluated as :

$$A = \begin{bmatrix} 0.1303 & -0.1303 & 0 \\ 0 & 0.6066 & -0.6066 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = 0$$

Replacing values of three, four, five and six state model and severity probabilities in equation (8.75) the aggregate loss probabilities are:

States	Actual	Geo-Wei	Geo-GP	Geo-Par	Geo-OPPL	Geo-TPPL
3 state	0.0002	$2.17 * e^{-19}$	$2.71 * e^{-20}$	$2.17 * e^{-19}$	$1.26 * e^{-29}$	$3.10 * e^{-25}$
4 state	0.0011	$5.29 * e^{-23}$	$5.29 * e^{-23}$	$5.29 * e^{-23}$	$6.16 * e^{-33}$	$5.05 * e^{-29}$
5 state	0.0014	0	0	0	$7.88 * e^{-31}$	0
6 state	0.0035	$5.29 * e^{-23}$	$4.24 * e^{-22}$	$2.64 * e^{-23}$	$7.70 * e^{-34}$	$6.31 * e^{-30}$

Table 8.19. Aggregate loss probabilities of phase type Geometric and severity distributions

Table (8.19) illustrate that the trend of actual aggregate losses are different from aggregate losses estimated using PH Geometric.

8.9.4 Aggregate Loss probabilities of phase type Binomial and severity distributions

The initial condition, matrices A and B of phase type Binomial distribution are calculated as:

$$P_0 = [I - P]^{\alpha} \qquad A = -P[I - P]^{-1} \qquad B = (\alpha + 1)P[I - P]^{-1}$$

Phase type Panjer recursive model when claim frequency distribution is phase type Binomial distribution is expressed as:

$$Z(j) = \sum_{x=1}^{j} y(x) Z(j - x) \left(-P[I - P]^{-1} + (\alpha + 1)P[I - P]^{-1} \frac{x}{j} \right) \left(I + P[I - P]^{-1} y(0) \right)^{-1} \qquad (8.76)$$

For three state multi-state model the initial condition is evaluated as:

$$P_0 = \begin{bmatrix} 0.0002 & 0.0002 & 0 \\ 0 & 0.1350 & 0.1350 \\ 0 & 0 & 0 \end{bmatrix} \Upsilon P_0 = \begin{bmatrix} 0 & 0.0074 & 0.0060 \end{bmatrix}$$

The value of $A = -P[I - P]^{-1}$ and value $B = (\alpha + 1)P[I - P]^{-1}$ hence for three state they can be evaluated as :

$$A = \begin{bmatrix} -7.2169 & 1.0000 & 0 \\ 0 & -0.6496 & 1.0000 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = 1.0 * e^{3} \begin{bmatrix} 3.3992 & -0.4710 & 0 \\ 0 & 0.3060 & -0.4710 \\ 0 & 0 & 0 \end{bmatrix}$$

Replacing values of three, four, five and six state model and severity probabilities in equation (8.76) the aggregate loss probabilities are:

States	Actual	Bin-Wei	Bin-GP	Bin-Par	Bin-OPPL	Bin-TPPL
3 state	0.0002	0	0	$-5.28 * e^{-102}$	$-1.08 * e^{-113}$	$-1.17 * e^{-109}$
4 state	0.0011	0	0	0	0	0
5 state	0.0014	0	$4.24 * e^{-97}$	$1.12 * e^{-100}$	$1.65 * e^{-113}$	$7.08 * e^{-110}$
6 state	0.0035	0	0	0	0	0

Table 8.20. Aggregate loss probabilities of phase type Binomial and severity distributions

Table (8.20) shows that the trend of actual aggregate losses are different from aggregate losses estimated using PH Binomial. Table (8.21) represents p-values of aggregate loss probabilities shown in table (8.17), table (8.18) and table (8.19) except for phase type Binomial model which had negative values.

Hypothesis	P-value	Multiple R
H_0 : Estimated aggregate losses for Panjer class $(a, b, 0)$ are not correlated to actual aggregate losses		
PH-Poisson Weibull	0.1276	0.7701
PH-Poisson Pareto	0.0927	0.8151
PH-Poisson Gen Pareto	0.0249	0.9239
PH-Poisson OPPL	0.4072	0.4854
PH-Poisson TPPL	0.4952	0.4080
PH-Neg Binom Weibull	0.9342	0.0517
PH-Neg Binom Pareto	0.8859	0.0898
PH-Neg Binom Gen Pareto	0.5525	0.3593
PH-Neg Binom OPPL	0.8963	0.0815
PH-Neg Binom TPPL	0.8963	0.0815
PH-Geo Weibull	0.9344	0.0515
PH-Geo Pareto	0.9346	0.0514
PH-Geo Gen Pareto	0.9164	0.0657
PH-Geo OPPL	0.9067	0.0733
PH-Geo TPPL	0.9347	0.0513

Table 8.21. P-values and Multiple R for Panjer class $(a, b, 0)$ aggregate loss models

Table (8.21) illustrates that the p-values for the models considered were greater than 0.05 except for PH-Poisson Generalized Pareto model which shows that this model provided a good fit for the data. The aggregate loss probabilities for class $(a, b, 0)$ and actual can be illustrated as:

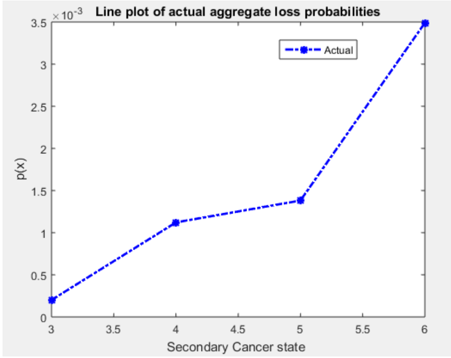
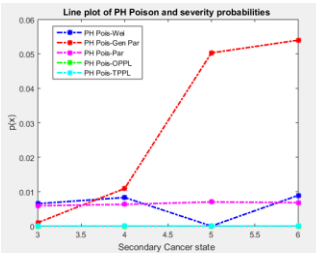
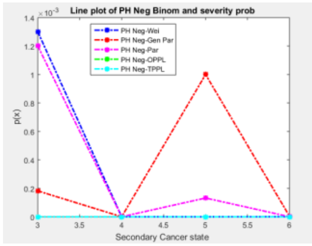


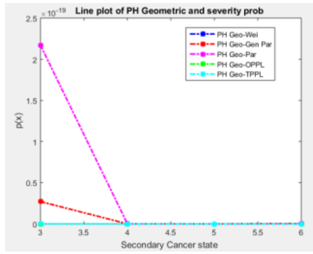
Figure 8.12. Aggregate loss probabilities for Actual data



(a) Aggregate Loss Probabilities PH Poisson Severity probabilities



(b) Aggregate Loss Probabilities PH Negative Binomial Severity probabilities



(c) Aggregate Loss Probabilities PH Geometric Severity probabilities

Figure 8.13. Comparison of Aggregate loss probabilities using PH Poisson, PH Negative Binomial and PH Geometric distributions

The data of aggregate loss probabilities using PH Panjer class ($a, b, 0$) can be represented graphically as shown in figure (8.13). Aggregate losses using PH Binomial distribution with the severity probabilities had numerous negative values hence they were not ploNed. Aggregate losses using PH Poisson and Generalized Pareto are similar to actual aggregate losses hence they are ploNed for comparison.

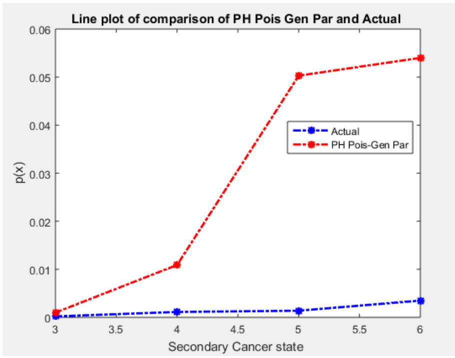


Figure 8.14. Comparison of Agg loss prob using PH Pois with Gen Par and Actual probabilities

Figure (8.14) shows comparison of aggregate loss probabilities using PH Poisson with Generalized Pareto and actual aggregate loss probabilities. PH Poisson and Generalized Pareto provided the best fit for class ($a, b, 0$). This is also supported by its p-value as it has the lowest p-value and the highest Multiple R.

8.10 Estimation Aggregate loss probabilities using Panjer recursive model for class $(a, b, 1)$

Phase type Panjer recursive model for class $(a, b, 1)$ is expressed as:

$$\vec{Z}(j) = [P_1 - (A+B)P_0]y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(A + B\frac{x}{j}\right)[I - Ay(0)]^{-1} \quad \textbf{(8.77)}$$

where $\vec{Z}(0) = \Upsilon P_1^T$ because P_1^T is the initial condition for Panjer class $(a, b, 1)$.

8.10.1 Aggregate Loss probabilities of phase type Zero -truncated Poisson and severity distribution

The initial condition, matrices A and B of phase type Zero-truncated Poisson distribution are calculated as:

$$P_1^T = P_1[I - P_0]^{-1} \qquad A = 0 \qquad B = \Lambda$$

Phase type Panjer recursive model when claim frequency distribution is Phase type Zero truncated Poisson distribution is expressed as:

$$\vec{Z}(j) = \Upsilon[P_1 - \Lambda P_0]y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(\Lambda\frac{x}{j}\right) \quad \textbf{(8.78)}$$

For three state multi-state model the initial condition is evaluated as:

$$P_1 = \begin{bmatrix} 0.3649 & 0.0249 & -0.0220 \\ 0 & 0.2656 & 0.1023 \\ 0 & 0 & 0.3679 \end{bmatrix} \qquad \Upsilon P_1 = \begin{bmatrix} 0.0876 & 0.0220 & 0.2583 \end{bmatrix}$$

The value of $A = 0$ and value $B = \Lambda$ hence for three state they can be evaluated as :

$$A = 0 \qquad B = \begin{bmatrix} 0.8788 & 0.1212 & 0 \\ 0 & 0.3981 & 0.6019 \\ 0 & 0 & 1.0000 \end{bmatrix}$$

Replacing values of three, four, five and six state model and severity probabilities in equation (8.78) the aggregate loss probabilities are:

States	Actual	ZT Poi-Wei	ZT Poi-GP	ZT Poi-Par	ZT Poi-OPPL	ZT Poi-TPPL
3 state	0.0002	0.0103	0.0016	0.0094	$9.37 * e^{13}$	$1.02 * e^{-8}$
4 state	0.0011	0.0131	0.0172	0.01	$1.19 * e^{12}$	$1.05 * e^{-8}$
5 state	0.0014	0	0.0796	0.0111	$7.79 * e^{-13}$	$3.34 * e^{-9}$
6 state	0.0035	0.0141	0.0855	0.0107	$1.89 * e^{-13}$	$1.63 * e^{-9}$

Table 8.22. Aggregate loss probabilities of phase type ZT Poisson and severity distributions

Table (8.22) shows that actual aggregate loss probabilities are similar to aggregate loss probabilities using Zero-truncated Poisson and Generalized Pareto.

8.10.2 Aggregate Loss probabilities of phase type Zero -truncated Binomial and severity distribution

The initial condition, matrices A and B of phase type Zero-truncated Binomial distribution are calculated as:

$$P_1^T = P_1[I - P_0]^{-1} \qquad A = -P[I-P]^{-1} \qquad B = (\alpha+1)P[I-P]^{-1}$$

Phase type Panjer recursive model when claim frequency distribution is Phase type Zero truncated Binomial distribution is expressed as:

$$\vec{Z}(j) = \Upsilon P_1 - \left(-P[I-P]^{-1} + (\alpha+1)P[I-P]^{-1}\right) y(j) + \sum_{x=1}^{j} y(x)\vec{Z}(j-x)\left(-P[I-P]^{-1} + (\alpha+1) P[I-P]^{-1}\frac{x}{j}\right)\left(I + P[I-P]^{-1} y(0)\right)^{-1} \quad \textbf{(8.79)}$$

For three state multi-state model the initial condition is evaluated as:

$$P_1 = 1.0e-99 * \begin{bmatrix} 0 & -0.3174 & -0.3174 \\ 0 & 0.2062 & 0.2062 \\ 0 & 0 & 0 \end{bmatrix} \qquad \Upsilon P_1 = 1.0e-100 * \begin{bmatrix} 0 & -0.6374 & -0.6374 \end{bmatrix}$$

The value of $A = -P[I-P]^{-1}$ and value $B = (\alpha+1)P[I-P]^{-1}$ hence for three state they can be evaluated as :

$$A = \begin{bmatrix} -7.2169 & 1.0000 & 0 \\ 0 & -0.6496 & 1.0000 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 3399.2 & -471 & 0 \\ 0 & 306 & -471 \\ 0 & 0 & 0 \end{bmatrix}$$

Replacing values of three, four, five and six state model and severity probabilities in equation (8.79) the aggregate loss probabilities are:

States	Actual	ZT Bin-Wei	ZT Bin-GP	ZT Bin-Par	ZT Bin-OPPL	ZT Bin-TPPL
3 state	0.0002	$7.6 * e^{-99}$	$5.9 * e^{-101}$	$5.4 * e^{-99}$	$1.7 * e^{-110}$	$1.8 * e^{-106}$
4 state	0.0011	0	0	0	0	0
5 state	0.0014	0	$-4.4 * e^{-94}$	$-9.5 * e^{-98}$	$-1.7 * e^{-110}$	$-7.1 * e^{-107}$
6 state	0.0035	0	0	0	0	0

Table 8.23. Aggregate loss probabilities of phase type ZT Binomial and severity distributions

Table (8.23) shows that actual aggregate loss probabilities are not similar to aggregate loss probabilities using Zero-truncated Binomial and severity distribution.

8.10.3 Aggregate Loss probabilities of phase type Zero -truncated Geometric and severity distribution

The initial condition, matrices A and B of phase type Zero-truncated Poisson distribution are calculated as:

$$P_1^T = P_1[I - P_0]^{-1} \qquad A = [I - P] \qquad B = 0$$

Phase type Panjer recursive model when claim frequency distribution is Phase type Zero truncated Geometric distribution is expressed as:

$$\bar{Z}(j) = \Upsilon P_1 [I - P] P_0 y(j) + \sum_{x=1}^{j} y(x) \bar{Z}(j-x)[I - P][I - [I - P]y(0)]^{-1} \qquad \textbf{(8.80)}$$

For three state multi-state model the initial condition is evaluated as:

$$P_1 = \begin{bmatrix} 0.1069 & -0.0331 & -0.0738 \\ 0 & 0.2387 & -0.2387 \\ 0 & 0 & 0 \end{bmatrix} \qquad \Upsilon P_1 = \begin{bmatrix} 0.0257 & 0.0064 & -0.0321 \end{bmatrix}$$

The value of $A = [I - P]$ and value $B = 0$ hence for three state they can be evaluated as :

$$A = \begin{bmatrix} 0.1303 & -0.1303 & 0 \\ 0 & 0.6066 & -0.6066 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = 0$$

Replacing values of three, four, five and six state model and severity probabilities in equation (8.80) the aggregate loss probabilities are:

States	Actual	ZT Geo-Weib	ZT Geo-GP	ZT Geo-Par	ZT Geo-OPPL	ZT Geo-TPPL
3 state	0.0002	$2.2 * e^{-19}$	$5.4 * e^{-20}$	$2.2 * e^{-19}$	$1.3 * e^{-29}$	$4.1 * e^{-25}$
4 state	0.0011	$5.3 * e^{-23}$	$5.3 * e^{-23}$	$5.3 * e^{-23}$	$6.2 * e^{-33}$	$5.1 * e^{-29}$
5 state	0.0014	0	0	$6.8 * e^{-21}$	$3.9 * e^{-31}$	0
6 state	0.0035	$5.3 * e^{-23}$	$4.2 * e^{-22}$	$2.6 * e^{-23}$	$7.7e^{-34}$	$6.3 * e^{-30}$

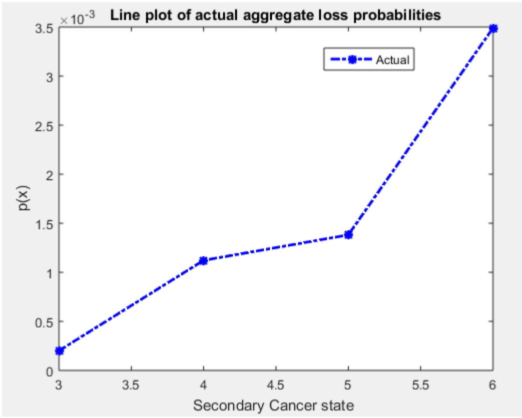
Table 8.24. Aggregate loss probabilities of phase type ZT Geometric and severity distributions

Table (8.24) shows that actual aggregate loss probabilities are not similar to aggregate loss probabilities using Zero-truncated Geometric and severity distribution. Table (8.25) represents p-values of aggregate loss probabilities shown in table (8.22) and table (8.24).

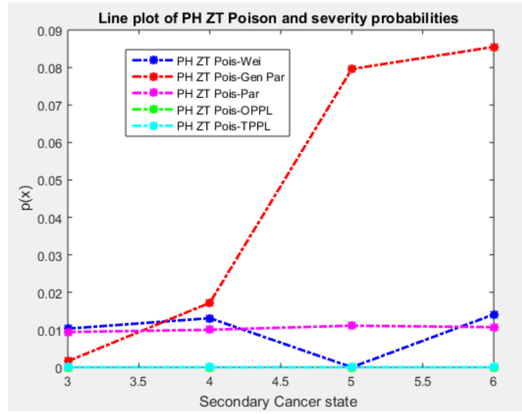
Hypothesis	P-value	Multiple R
H_0 : Estimated aggregate losses for Panjer class $(a, b, 1)$ are not correlated to actual aggregate losses		
PH-ZT Poisson Weibull	0.1273	0.7705
PH-ZT Poisson Pareto	0.0946	0.8125
PH-ZT Poisson Gen Pareto	0.0249	0.9239
PH-ZT Poisson OPPL	0.4971	0.4854
PH-ZT Poisson TPPL	0.4952	0.4080
PH-ZT Geometric Weibull	0.9344	0.0515
PH-ZT Geometric Pareto	0.9206	0.0624
PH-ZT Geometric Gen Pareto	0.9255	0.0585
PH-ZT Geometric OPPL	0.9205	0.0624
PH-ZT Geometric TPPL	0.9347	0.0513

Table 8.25. P-values and Multiple R for Panjer class $(a, b, 1)$ aggregate loss models

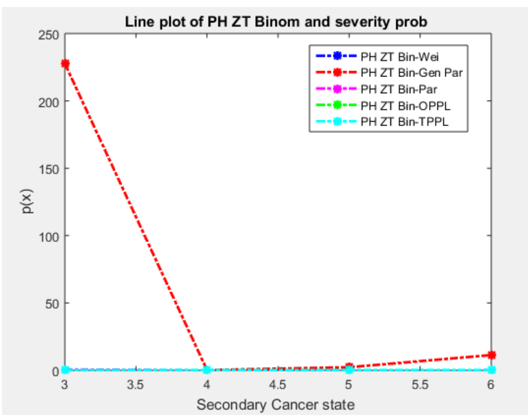
Table (8.25) indicates that the p-values for the models considered were greater than 0.05 except for PH ZT Poisson Generalized Pareto, indicating that it was the model which provided the good fit for the data. The data of aggregate loss probabilities using PH Panjer class (a, b, 1) and actual data can be represented graphically as shown in figure (8.15) as:



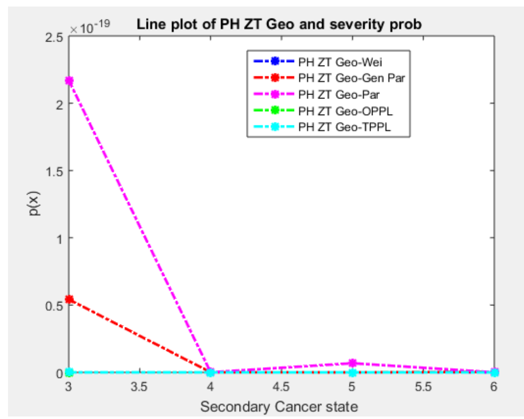
(a) Actual Aggregate Loss Probabilities



(b) Aggregate Loss Probabilities PH ZT Poisson Severity probabilities



(c) Aggregate Loss Probabilities PH ZT Binomial Severity probabilities



(d) Aggregate Loss Probabilities PH ZT Geometric Severity probabilities

Figure 8.15. Comparison of Aggregate loss probabilities using PH ZT Poisson, PH ZT Binomial and PH ZT Geometric distributions

Figure (8.15) shows comparison of aggregate loss probabilities using PH ZT Poisson, PH ZT Binomial, PH ZT Geometric distributions with severity probabilities. It is evident that aggregate loss probabilities using PH ZT Poisson Generalized Pareto model were similar to actual probabilities hence they are ploNed for comparison. PH ZT Poisson Generalized Pareto had the lowest p-values and the highest Multiple R values hence it was considered as the best model. Figure (8.16) shows comparison of aggregate loss probabilities for PH ZT Poisson Generalized Pareto and actual probabilities.

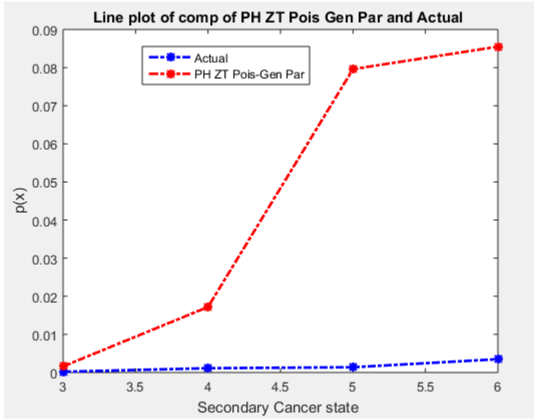


Figure 8.16. Comparison of Agg loss prob PH ZT Pois-Gen Par and Actual

Figure (8.16) shows that PH ZT Poisson Generalized Pareto provided a good fit although it slightly over estimated the aggregate loss probabilities. The model that had the lowest p-value and highest Multiple R was PH ZT Poisson Generalized Pareto model hence it was considered the best model. Figure (8.17) shows comparison of aggregate loss probabilities using best model for Panjer class $(a, b, 0)$ and best model for Panjer class $(a, b, 1)$.

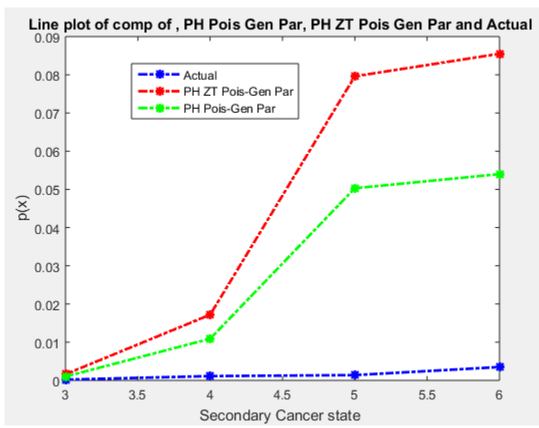


Figure 8.17. Comparison of Agg loss prob Pois- Gen Par , PH ZT Pois-Weib and Actual

Figure (8.17) shows that PH Poisson Generalized Pareto model provided a beNer fit compared to PH ZT Poisson Generalized Pareto.

8.11 Estimation Aggregate loss probabilities using Poisson Lindley distributions

Discrete Fourier Transforms will be used to estimate aggregate loss probabilities of Poisson Lindley distributions. Discrete Fourier Transform is expressed as:

$$X_k = \sum_{n=0}^{N-1} X(n) e^{\frac{-i2\Pi kn}{N}} \qquad k = 0, 1, 2, \ldots, N-1 \tag{8.81}$$

Expression (8.81) is complex to compute hence it is simplified using Euler's formula to:

$$X(k) = \sum_{n=0}^{N-1} X(n) \cos \frac{2\Pi kn}{N} - i \sin \frac{2\Pi kn}{N} \qquad X(k) = \sum_{n=0}^{N-1} X(n) W_N^{kn} \tag{8.82}$$

which is the expression of DFT of claim amount or claim count probabilities. The claim amount and claim count probabilities are of the length 8, hence the matrix W is a primitive 8^{th} root of unity leading to equation (8.82) being:

$$X(k) = \sum_{N=0}^{7} X(n) W_N^{kn} \tag{8.83}$$

8.11.1 Aggregate Loss probabilities of phase type one and two parameter Poisson Lindley and severity distribution

The DFT of frequency is calculated by multiplying the matrix W_N^{kn} and phase type one parameter Poisson Lindley and phase type two parameter Poisson Lindley distributions which is expressed mathematically as:

$$\text{DFT of frequency} = W_N^{kn} * \text{frequency /severity probabilities}$$

where in this case the frequency probabilities belong to phase type one parameter Poisson Lindley and phase type two parameter Poisson Lindley distributions and severity probabilities belong to Weibull, Generalized Pareto, Pareto, one parameter Poisson Lindley and two parameter Poisson Lindley distributions.

8.11.2 Tabulation of aggregate loss probabilities using OPPL and TPPL distributions

The aggregate loss probabilities using phase type one parameter Poisson Lindley distribution and severity distributions as obtained from equation (8.83) are:

States	Actual	PH-OPPL Wei	PH-OPPL Par	PH-OPPL GP	PH-OPPL OPPL	PH-OPPL TPPL
3 state	0.0002	0.00028	0.00026	0.00045	$2.6971 * e^{-14}$	$2.9146 * e^{-10}$
4 state	0.0011	0.00057	0.00047	0.0005	$5.4223 * e^{-14}$	$5.1986 * e^{-10}$
5 state	0.0014	0.00044	0.00067	0.00223	$6.3975 * e^{-14}$	$4.9640 * e^{-10}$
6 state	0.0035	0.00074	0.00083	0.00365	$5.5951 * e^{-14}$	$4.4436 * e^{-10}$

Table 8.26. Aggregate loss probabilities of phase type OPPL and severity distributions

Table (8.26) illustrates that phase type one parameter Poison Lindley with Weibull, Pareto and Generalized Pareto are similar to actual aggregate loss probabilities. The aggregate loss probabilities using phase type two parameter Poisson Lindley distribution and severity distributions as obtained from equation (8.83) are:

States	Actual	PH-TPPL Wei	PH-TPPL Par	PH-TPPL GP	PH-TPPL OPPL	PH-TPPL TPPL
3 state	0.0002	0.00024	0.00022	0.00003	$2.2921 * e^{-14}$	$2.4783 * e^{-10}$
4 state	0.0011	0.00051	0.00042	0.00043	$4.7997 * e^{-14}$	$4.6248 * e^{-10}$
5 state	0.0014	0.00042	0.00061	0.00193	$5.85556 * e^{-14}$	$4.6240 * e^{-10}$
6 state	0.0035	0.00067	0.00075	0.00327	$5.2901 * e^{-14}$	$4.2082 * e^{-10}$

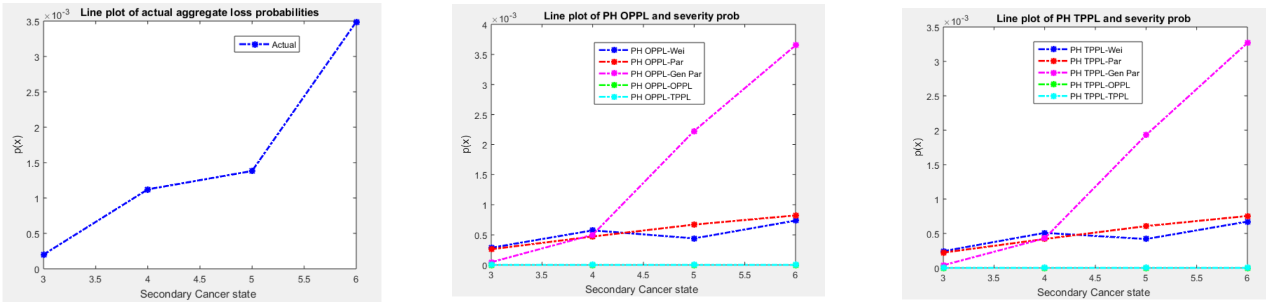
Table 8.27. Aggregate loss probabilities of phase type TPPL and severity distributions

Table (8.27) illustrates that phase type OPPL with Weibull, Pareto and Generalized Pareto are similar to actual aggregate loss probabilities.

Hypothesis	P-value	Multiple R
H_0 : Estimated aggregate losses for Poisson Lindely are not correlated to actual aggregate losses		
PH OPPL-Weibull	0.0243	0.9250
PH OPPL-Pareto	0.0184	0.9378
PH OPPL-Gen Pareto	0.0061	0.9703
PH OPPL-OPPL	0.0649	0.8547
PH OPPL-TPPL	0.0885	0.8208
PH TPPL-Weibull	0.0222	0.9294
PH TPPL-Pareto	0.0164	0.9424
PH TPPL-Gen Pareto	0.0056	0.9717
PH TPPL-OPPL	0.0564	0.8678
PH TPPL-TPPL	0.0778	0.8358

Table 8.28. P-values and Multiple R for Poisson Lindley aggregate loss models

Table (8.28) illustrates that the p-values for the models considered were less than 0.05 for PH OPPL with continuous severity distributions and PH TPPL with continuous severity distributions indicating that they were the significant models for mixture models. Aggregate loss probabilities in table (8.26) and (8.27) can be represented graphically as:



(a) Actual aggregate loss probabilities

(b) Aggregate Loss Probabilities PH OPPL Severity probabilities

(c) Aggregate Loss Probabilities PH TPPL Severity probabilities

Figure 8.18. Comparison of Aggregate loss probabilities using PH OPPL and PH TPPL distributions

Figure (8.18) shows that PH OPPL Generalized Pareto and PH TPPL Generalized Pareto models were similar to actual aggregate loss probabilities hence they will compared to determine which provides a beNer fit. Figure (8.19) shows comparison of aggregate loss probabilities using PH OPPL with Generalized Pareto, PH TPPL with Generalized Pareto and actual probabilities.

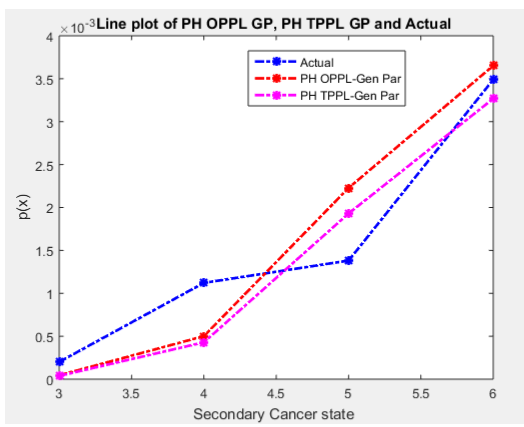


Figure 8.19. Comparison of Agg loss pro PH OPPL-Gen Par, PH TPPL-Wei, PH TPPL-Gen Par and Actual

Figure (8.19) shows that PH TPPL with Generalized Pareto provided a beNer fit compared to PH OPPL with Generalized Pareto model. This is also supported by its p-values and Multiple R values which were the lowest compared to all the other models.

8.11.3 Aggregate Loss probabilities of phase type ZT one and ZT two parameter Poisson Lindley and severity distribution

The DFT of frequency is calculated by multiplying the matrix W_N^{kn} and phase type Zero-truncated one parameter Poisson Lindley and phase type Zero truncated two parameter Poisson Lindley distributions which is expressed mathematically as:

$$\text{DFT of frequency} = W_N^{kn} * \text{frequency /severity probabilities}$$

where in this case the frequency probabilities belong to phase type Zero truncated one parameter Poisson Lindley and phase type Zero truncated two parameter Poisson Lindley distributions and severity probabilities belong to Weibull, Generalized Pareto, Pareto, one parameter Poisson Lindley and two parameter Poisson Lindley distributions.

8.11.4 Tabulation of aggregate loss probabilities using ZT OPPL and ZT TPPL distributions

The aggregate loss probabilities using phase type Zero-truncated one parameter Poisson Lindley distribution and severity distributions as obtained from equation (8.83) are:

States	Actual	PH-ZT OPPL Wei	PH-ZT OPPL Par	PH-ZT OPPL GP	PH-ZT OPPL OPPL	PH-OPPL TPPL
3 state	0.0002	0.00046	0.00042	0.00007	$4.3099 * e^{-14}$	$4.6599 * e^{-10}$
4 state	0.0011	0.00092	0.00076	0.00080	$8.6717 * e^{-14}$	$8.3142 * e^{-10}$
5 state	0.0014	0.00071	0.00107	0.00356	$1.0233 * e^{-13}$	$7.9414 * e^{-10}$
6 state	0.0035	0.00118	0.00131	0.00583	$8.9547 * e^{-14}$	$7.1121 * e^{-10}$

Table 8.29. Aggregate loss probabilities of phase type ZT OPPL and severity distributions

Table (8.29) shows that phase type Zero-truncated one parameter Poisson Lindley with Weibull, Pareto and Generalized Pareto are similar to actual aggregate loss probabilities. The aggregate loss probabilities using phase type Zero-truncated two parameter Poisson Lindley distribution and severity distributions as obtained from equation (8.83) are:

States	Actual	PH-ZT TPPL Wei	PH-ZT TPPL Par	PH-ZT TPPL GP	PH-ZT TPPL OPPL	PH-ZT TPPL TPPL
3 state	0.0002	0.00026	0.00024	0.00004	$2.4535 * e^{-14}$	$2.6528 * e^{-10}$
4 state	0.0011	0.00054	0.00044	0.00046	$5.1165 * e^{-14}$	$4.9276 * e^{-10}$
5 state	0.0014	0.00044	0.00064	0.00206	$6.1967 * e^{-14}$	$4.8778 * e^{-10}$
6 state	0.0035	0.00071	0.00079	0.00347	$5.5426 * e^{-14}$	$4.4011 * e^{-10}$

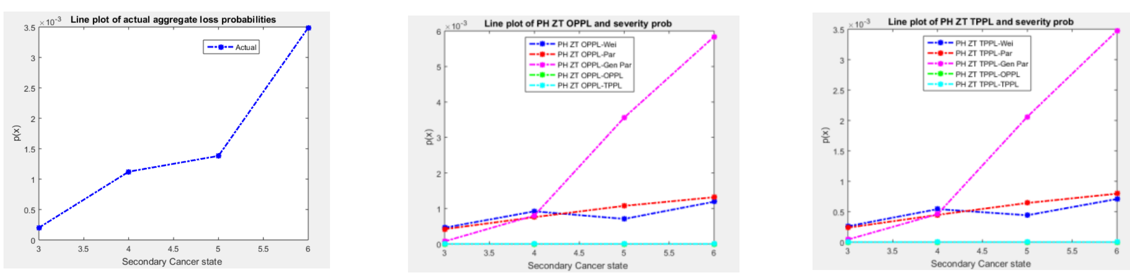
Table 8.30. Aggregate loss probabilities of phase type ZT TPPL and severity distributions

Table (8.30) shows that phase type OPPL with Weibull, Pareto and Generalized Pareto are similar to actual aggregate loss probabilities.

Hypothesis	P-value	Multiple R
H_0 : Estimated aggregate losses for ZT Poisson Lindely are not correlated to actual aggregate losses		
PH ZT OPPL-Weibull	0.0243	0.9251
PH ZT OPPL-Pareto	0.0184	0.9378
PH ZT OPPL-Gen Pareto	0.0061	0.9702
PH ZT OPPL-OPPL	0.0648	0.8549
PH ZT OPPL-TPPL	0.0883	0.8211
PH ZT TPPL-Weibull	0.0230	0.9277
PH ZT TPPL-Pareto	0.0168	0.9414
PH ZT TPPL-Gen Pareto	0.0057	0.9715
PH ZT TPPL-OPPL	0.0584	0.8647
PH ZT TPPL-TPPL	0.0806	0.8318

Table 8.31. P-values and Multiple R for ZT Poisson Lindley aggregate loss models

Table (8.31) illustrates that the p-values for the models considered were less than 0.05 for PH ZT OPPL with continuous severity distributions and PH ZT TPPL with continuous severity distributions indicating that they were the significant models. Aggregate loss probabilities in table (8.29) and (8.30) can be represented graphically as:



(a) Actual aggregate loss probabilities

(b) Aggregate Loss Probabilities PH ZT OPPL Severity probabilities

(c) Aggregate Loss Probabilities PH ZT TPPL Severity probabilities

Figure 8.20. Comparison of Aggregate loss probabilities using PH ZT OPPL and PH ZT TPPL distributions

Figure (8.20) shows that PH ZT OPPL with Generalized Pareto and PH ZT TPPL with Generalized Pareto were similar to actual aggregate loss probabilities hence they will compared to determine which provides a beNer fit. Figure (8.21) shows comparison of aggregate loss probabilities using PH ZT OPPL with Generalized Pareto and PH ZT TPPL with Generalized Pareto and actual probabilities.

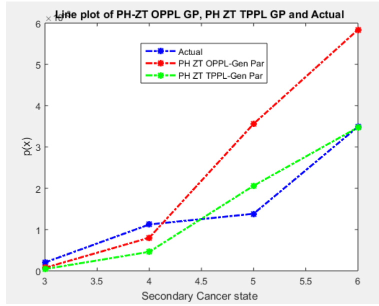


Figure 8.21. Comparison of Agg loss pro PH ZT OPPL-Gen Par, PH ZT TPPL-Gen Par and Actual

Figure (8.21) shows that PH-ZT TPPL Generalized Pareto model provided a beNer fit compared to PH-ZT OPPL Generalized Pareto model. Figure (8.22) shows comparison of aggregate loss probabilities using PH TPPL with Generalized Pareto and PH ZT TPPL with Generalized Pareto and actual probabilities.

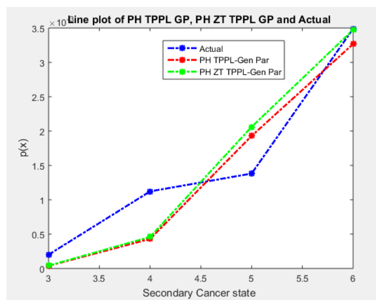


Figure 8.22. Comparison of Agg loss pro PH TPPL-Gen Par, PH ZT TPPL-Gen Par and Actual

Figure (8.22) shows both PH TPPL Generalized Pareto and PH ZT TPPL Generalized Pareto were very close to be differentiated graphically. They can be chosen using p-value and Multiple R values where PH TPPL Generalized Pareto model is superior. Figure (8.23) shows comparison of aggregate loss probabilities using PH Poisson Generalized Pareto, PH TPPL Generalized Pareto models and actual probabilities.

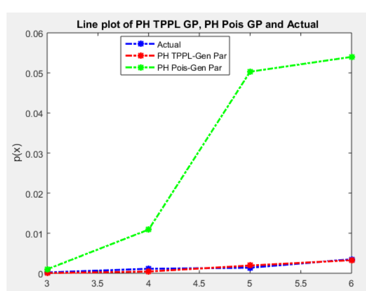


Figure 8.23. Comparison of Agg loss pro PH Pois-Gen Par, PH TPPL-Gen Par and Actual

Figure (8.23) shows that PH TPPL Generalized Pareto model provided the overall best model for secondary cancer data. This model had also the lowest p-values and Multiple R.

9 CONCLUSIONS AND RECOMMENDATIONS

9.1 Summary of Results and Challenges

The objective of this study is to estimate aggregate loss probabilities using Phase type distributions. Phase type distributions of class $(a, b, 0)$, class $(a, b, 1)$ and phase type Poisson Linldey distributions were considered as frequency distributions and both continuous and discrete distributions were considered in modeling severity data. Chapman-Kolmogorov equations were considered in construction of multi-state models. This research considered three, four, five and six state secondary cancer multi-state models. Phase type distributions improves estimation of aggregate loss probabilities of diseases which are dynamic like cancer. The best phase type distribution of class $(a, b, 0)$, class $(a, b, 1)$ and phase type Poisson Lindley distributions in estimation of aggregate loss probabilities were determined.

9.1.1 Frequency and Severity distributions.

The claim frequency distributions were phase type distributions from Panjer class $(a, b, 0)$, class $(a, b, 1)$ and phase type Poisson Lindley distributions. Phase type distribution of class $(a, b, 0)$ considered were PH Poisson, PH Negative Binomial , PH Binomial and PH Geometric distributions, for class $(a, b, 1)$ the distributions considered were PH Zero Truncated Poisson, PH Zero truncated Binomial and PH Zero truncated Geometric distributions and PH one parameter Poisson Lindley, PH two parameter Poisson Lindley, PH ZT one parameter Poisson Lindley and PH ZT two parameter Poisson Lindley for Poisson Lindley distributions. The frequency probabilities using PH distributions of Panjer class $(a, b, 0)$ are shown in figure (8.2) and showed that PH Poisson had a similar trend as the actual frequency probabilities. PH Poisson was compared with actual probabilities as shown in figure (8.3) and PH Poisson modeled the frequency data best for class $(a, b, 0)$. Phase type Poisson improves estimation of frequency distribution by in cooperating the transition probability of secondary cancers which can not be achieved using ordinary Poisson distribution. Phase type Poisson distribution solved the problem faced by insurance firms of underestimation of claim frequency which consequently results to errors in estimation of aggregate losses. This can lead insurance firms in to ruin as the reserves set aside will not be able to cater for the claims.

Frequency probabilities using PH ZT Poisson decreased from three state model to six state model. PH ZT Binomial probabilities had probabilities of zero for four state and six state model. PH ZT Geometric probabilities had zero probabilities for four state and five state models. PH ZT Poisson had frequency probabilities which were similar to actual data as shown in figure (8.4). Comparing PH ZT Poisson and actual probabilities as shown in figure (8.5) PH ZT Poisson provided the best fit for frequency data using distribution of PH Panjer class $(a, b, 1)$. PH OPPL and TPPL distributions are as shown in figure (8.6) and it evident that they are similar to actual frequency probabilities hence were compared to determined which provided the best fit. Figure (8.7) shows that phase type TPPL provided a good fit for all the states. PH ZT OPPL and ZT TPPL distributions are as shown in figure (8.8) and it evident that they are similar to actual frequency probabilities hence were compared to determined which provided the best fit. Figure (8.9) compares all the best frequency models but it was differentiate the best model hence P values and R statistics are used. The two best distributions for class $(a, b, 0)$ and $(a, b, 1)$ are compared in figure (8.9) and shows that PH TPPL provided the best overall model for the two classes of distributions based on Multiple R value.

Severity probabilities using both discrete and continuous distributions are shown in figure (8.10). The severity probabilities of Generalized Pareto distribution increased between the three state model and six state model which is similar to the actual data. Figure (8.11) compares Generalized Pareto and actual probabilities and shows that Generalized Pareto provides a good fit for severity probabilities. One parameter Poisson Lindley distribution and two parameter Poisson Lindley distribution were totally different from the actual data hence they can not be considered as good models for cancer severity data as they under estimate the severity probabilities which gives a negative effect on the insurance companies.

9.1.2 Panjer recursive model

Estimation of aggregate loss distributions using Phase type Panjer recursive method depend mainly on the primary distribution which is the frequency distribution. Aggregate loss probabilities of PH Panjer distributions of class $(a, b, 0)$ are shown in figure (8.13). Aggregate loss probabilities using PH Poisson with Generalized Pareto provided a good fit for secondary cancer data. The aggregate loss probabilities estimated using PH Poisson distribution with Generalized Pareto distribution had the same trend as actual aggregate loss probabilities hence it is considered as a good model for modeling aggregate losses for secondary cancer data. Comparing PH Poisson Generalized Pareto model with actual aggregate loss probabilities as illustrated in figure (8.14) PH Poisson-Generalized Pareto provided the best fit as it was a good model for all state although it slightly overestimated aggregate loss probabilities. PH Poisson-Weibull, Pareto, OPPL and TPPL model had aggregate losses probabilities lower than the actual aggregate probabilities hence they did not model the effect of in-cooperating the transition probabilities hence they can not be considered as a good model for secondary cancer cases. PH Negative Binomial and PH Geometric model had aggregate loss probabilities which were way lower than the actual aggregate loss probabilities hence they can not be considered as a good fit as it underestimate the aggregate losses. This research concludes that PH Poisson with Generalized Pareto distribution provided the best model for modeling aggregate loss probabilities for class $(a, b, 0)$. PH Poisson with one parameter Poisson Lindley and two parameter Poisson Lindley under estimated the aggregate loss probabilities which could lead to under estimation of reserves which can lead to insurance firms going to ruin. Aggregate loss probabilities using PH Negative Binomial and severity reduced from three state to the four state model, increased to five state model and reduced for six state model which is contrary with actual aggregate loss probabilities hence PH Negative Binomial model was not a good model with any of the severity distribution for secondary cancer data.

Frequency probabilities using PH distributions of Panjer class $(a, b, 1)$ are shown in figure (8.15). Aggregate loss probabilities using PH ZT Poisson with Weibull, Pareto, OPPL and TPPL were low hence they did not provide a good fit for the aggregate loss probabilities. PH ZT Poisson with

Generalized Pareto distributions were similar to the actual aggregate loss distributions hence they are compared with the actual data as shown in figure (8.16) to determine if it provides a good fit for secondary cancer data. Comparing the PH ZT Poisson- Generalized Pareto model with actual data it shows that slightly overestimated the aggregate losses. Aggregate loss probabilities using PH ZT Geometric distribution with severity distributions and PH ZT Binomial distribution with severity distributions were inconsistent with real data hence they could not be considered as good models in modeling secondary cancer data. Aggregate loss probabilities using distributions of Panjer class $(a, b, 1)$ were best modeled using ZT Poisson distribution and Generalized Pareto distribution. Aggregate loss probabilities for Panjer class $(a,\ b,\ 0)$, PH Poisson and Generalized Pareto distribution provided the best fit while for Panjer class $(a,\ b,\ 1)$, the best model was PH ZT Poisson with Generalized Pareto distribution. The two models were compared as shown in figure (8.17) and PH Poisson with Generalized Pareto distribution provided the best fit for the data between the two models.

9.1.3 Phase type Poisson Lindley model

Aggregate loss probabilities of PH OPPL and PH TPPL with severity probabilities are shown in figure (8.18). Aggregate loss probabilities using PH OPPL with Generalized Pareto and PH TPPL with Generalized Pareto provided a good fit for secondary cancer data. The aggregate loss probabilities estimated using these two models had the same trend as actual aggregate loss probabilities hence they are compared to determine which model provide the best fit for secondary cancer cases. Comparing the two models with actual aggregate loss probabilities as illustrated in figure (8.19) indicated that PH TPPL-Generalized Pareto provided the best fit as it was a good model for all states. PH OPPL and PH TPPL with one parameter Poisson and two parameter Poisson under estimated the aggregate loss probabilities which could lead to under estimation of reserves which can consequently lead to insurance firms going into ruin.

Aggregate loss probabilities of PH ZT OPPL and PH ZT TPPL with severity probabilities are shown in figure (8.20). Aggregate loss probabilities using PH ZT OPPL with Generalized Pareto and PH ZT TPPL with Generalized Pareto provided a good fit for secondary cancer data. The aggregate loss probabilities estimated using these two models had the same trend as actual aggregate loss probabilities hence they are compared to determine which model provide the best fit for secondary cancer cases. Comparing the two models with actual aggregate loss probabilities as shown in figure (8.21) indicated that PH ZT TPPL-Generalized Pareto provided the best fit as it was a good model for all states although it slightly overestimated the aggregate loss probabilities for five state model. PH ZT OPPL and PH ZT TPPL with one parameter Poisson and two parameter Poisson under estimated the aggregate loss probabilities which could lead to under estimation of reserves which can consequently lead to insurance firms going into ruin.

Comparing the best models PH TPPL Generalized Pareto, PH ZT TPPL Generalized Pareto with actual aggregate loss probabilities as shown in figure (8.22) indicated that PH TPPL-Generalized Pareto provided the best fit as it was a good model for all states. Figure (8.23) compares the best models for PH Panjer classes and PH Poisson Lindley models with actual aggregate loss probabilities indicating that PH TPPL-Generalized Pareto provided the best fit as it was a good model for all states.

9.2 Recommendation

The following recommendations have been suggested from this research.

9.2.1 Frequency and Severity

More research can be done on other Phase type distributions in estimation of claim frequency distributions as well as severity distributions where it is applicable. Research can also be done on discretization of severity probabilities using the method of local moment matching.
The aggregate loss probabilities were only considered for secondary cancer data hence this method can be applied in other diseases which can apply multi-state model such as Hiv-Aids.

9.2.2 Phase type Panjer recursive method

Other distributions of phase type Panjer class $(a, b, 1)$ can be used to estimate the frequency probabilities.
Aggregate loss probabilities using different data set for other diseases can be modeled using phase type Panjer recursive model.

9.2.3 Poisson Lindley model

Other Poisson Lindley distributions such as the three parameter Poisson Lindley distributions can be used to estimate the frequency probabilities.
Aggregate loss probabilities using different data set for other diseases can be modeled using phase type Poisson Lindley models.

9.2.4 Policy implications

This research can be used to improve estimation of aggregate losses of diseases which have multiple transition states. Policies can be drafted on how to implement this research on current policies in a way that will be beneficial for both the insurance firms as well as the policyholders as it will greatly improve access of health care for these diseases. This research methodology can also be extended to other field which in-cooperate multi- state models.

Bibliography

[Abdulhakim et. al, 2021] Abdulhakim A. Al-Babtain , Ahmed M. Gemeay and Ahmed Z. Afify (2021) . Estimation methods for the discrete Poisson-Lindley and discrete Lindley distributions with actuarial measures and applications in medicine . *Journal of King Saud University* Vol. 33, 101224.

[Asger et. al, 2018] Asger H. Arno S. and Mogens B (2018). Phase- type distribution in population genetics. *hEps://www.researchgate.net/publication/325592530*

[Asmussen, 1996] Asmussen S. and Bladt M. (1996). Phase-type distribution and risk processes with state dependent premiums. *Scand Actuar .j.*, No.1 ,19-36.

[Surya, 2016] B.A. Surya(2016) . Generalized Phase-Type Distribution and Competing Risks for Markov Mixtures Process.*Journal of statistics* arXiv:1611.03832v1

[Christian, 2006] Christian Hipp (2006). Phase-type distribution and its application in insurance. *hEp://insurance fbv.uni-karisruheda, university of Kalisruhe.*

[Clement et. al, 2019] Clement Twumasi, Louis Asiedu and Ezekiel N. (2019). Markov Chain Modeling of HIV, Tuberculosis, and Hepatitis B Transmission in Ghana .*Interdisciplinary Perspectives on Infectious Diseases* hNps://doi.org/10.1155/2019/9362492, Article ID 9362492 pp. 1-8

[Dmitry et.al, 2001] Dmitry E.Papush,Aleksey S. Popelyukhin and Jasmine G. Zhang (2001). Approximating the Aggregate Loss Distribution.*CAS Forum*, pp. 175 - 186.

[Eisele, 2006] Eisele Karl-Theodor (2006). Recursions for compound phase distributions.*Journal of Insurance: Mathematics and Economics* Vol. 38, pp. 149–156 .

[Elsayed, 2011] Elsayed A. Habib Elamir (2011).Zero-truncated Multiplicative-Binomial Distribution.*Applied Mathematics Sciences*,Vol.5,No.62,pp. 3073-3087.

[Erlang, 1909] Erlang, A (1909). Sandsynlighedsregning og telefonsamtaler. *Nyt tidsskrifl for Matematik*, Vol. 20, pp. 33-39 DOI : 10.1017/cbo9780511546594.

[Ernst, 1951] Ernst Hjalmar Waloddi Weibull (1951) . Weibull distribution . *American Society of Mechanical Engineers (ASME)* ,hNps://en.wikipedia.org/wiki/Weibull distribution.

[Fackler, 2009] Fackler Michael (2009). Panjer class united-one formula for poisson; Binomial and Negative Binomial distribution. *ASTIN Colloqulum: International Actuarial Association*.

[Felix et. al, 2014] Felix Noyanim Nwobi and Chukwudi Anderson Ugomma (2014) .A Comparison of Methods for the Estimation of Weibull Distribution Parameters. *Metodološki zvezki* Vol. 11, No. 1, pp. 65-75.

[Gabor et. al, 2012] Gábor Horvath and Philipp Reinecke (2012). Phase-type Distributions for Realistic Modeling in Discrete-Event Simulation .*DOI: 10.4108/icst.simutools.2012.247727*

[Gareth et. al, 2015] Gareth W. Peters and Pavel V. Shevchenko (2015). Advances in Heavy Tailed Risk Modeling. *ISBN:9781118909539, Wiley*, hNps://onlinelibrary.wiley.com/doi/book/10.1002/9781118909560 .

[Glenn, 1981] Glenn G. Meyers (1981). Estimating Predictive Distribution for Loss Reserve.*Casualty Actuary Society*,Vol 01/ISSUE 02,pp. 248-272.

[Guilherme et. al, 2016] Guilherme Ferraz , Francisco A. , Pablo Martin , Emanuele Cozzo, and Yamir Moreno (2016) . Unifying Markov Chain Approach for Disease and Rumor Spreading in Complex Networks .*arXiv:1609.00682v2 [physics.soc-ph]* DOI: 10.1093/comnet/cnx024

[Harchol, 2012] Harchol-Balter, M (2012). Real-world Workloads: High Variability and Heavy Tails. *Design of computer systems* p.347. doi:1017/CB09781139226424.026.

[Harry, 1981] Harry.H.Panjer (1981). Recursive Evaluation of a Family of compound distribution.*ASTIN Bulletin: International Actuarial Association* Vol.12, No.1 ,22-26. doi:10.1017/50515036100006796.

[Heckman et. al, 1983] Heckman,P. E. and Meyers,G. G.(1983).The calculation of aggregate loss distributions from claim severity and claim count distributions.*In Proceedings of the Casualty Actuarial Society*,Vol. 70, No. 133-134.

[Hess et. al, 2002] Hess Klaus Th, Liewald AneN, Schmidt Klaus D. (2002). An extension of Panjer's recursion. *ASTIN Bulletin* Vol.32, No.2, 282-297.

[Hewit, 1967] Hewit C.C (1967). Loss Ratio Distribution:A model.*LIV*,pp.70-88.

[Jane et. al, 2015] Jane M. , Rebecca A. , Lurdes Y., and Vladimir N. (2015). A Joint Model for Multistate Disease Processes and Random Informative Observation Times, with Applications to Electronic Medical Records Data .*The International Biometric Society* Vol. 71, pp. 90-101.

[Jensen, 1953] Jensen, A. (1953). A distribution model applicable to economics.. *Munksqaard*,Copenhagen .

[John et. al, 2016] John G. Kemeny and J. Laurie Snell (2016) . Finite Markov chains . *Princeton, NJ: Van Nostrand* ISBN 0-387·90192-2 Springer-Verlag.

[Garrido, 2006] Jose Garrido(2006). Claim Number Processes. *Wiley*.

[Juan, 2020] Juan Eloy Ruiz-Castro and Mariangela Zenga (2020). A general piecewise multi-state survival model: application to breast cancer. *Statistical methods and Application, Springer*, vol. 29, No. 4, pp. 813-843 .

[Kehinde et. al, 2019] Kehinde Adiguna, Ayan Adelekeb, Oluwasesan Adewusic, Oluwafunmilola Olubiyid, Omobolaji Halide and Bayowa Babalolaf(2019) . Modelling Infectious Diseases Using Markov Chain .*American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)* , Vol. 61, No. 01, pp. 280-288

[Kishore et. al, 2018] Kishore K Das, Inzamul Ahmed and Sahana BhaNacharjee (2018). A New Three-Parameter Poisson-Lindley Distribution for Modelling Over-dispersed Count Data . *International Journal of Applied Engineering* Vol. 13, No. 23, pp. 16468-16477.

[Kok et. al, 2010] Kok Keng Siaw and Xueyuan Wu (2010). Matrix-form recursive evaluation of the aggregate claims distribution.

[Kumer et. al, 2011] Kumer Pial Das, Shamim Sarker and Norou Diawara (2011). Further Review of Panjer's Recursion for Evaluation of Compound Negative Binomial Distribution Using R. *MSC2010: 60G44, 62P05*, DOI: 10.35834/mjms/1321045146, hNps://www.researchgate.net/publication/254211580, Vol. 23, No. 2, pp 182-190.

[Manuel et. al, 2021] Manuel L. , Gracinda R. , Matilde C. and Pedro Corte (2021). Calibration of Transition Intensities for a Multistate Model: Application to Long-Term Care.*Journal of Risk*,Vol. 9,No. 37, pp. 1-17

[Marcelo et. al, 2015] Marcelo G. Cruz, Gareth W. Peters and Pavel V. Shevchenko(2015). Fundamental Aspects of Operational Risk and Insurance Analytics: A Handbook of Operational Risk.*Wiley, ISBN: 978-1-118-11839-9* hNps://www.wiley.com/en-us/Fundamental+Aspects+of+Operational+Risk+and+Insurance+Analytics3A+A+Handbook+of+Opera p-9781118118399 .

[Mogens et. al, 2017] Mogens Bladt and Bo Friis Nielsen (2017). Matrix-Exponential Distributions in Applied Probability.*Springer Science and Business Media LLC*, hNps://link.springer.com/book/10.1007/978-1-4939-7049-0 .

[Mogens, 2005] Mogens Bladt (2005). A review on phase-type distribution and their use in Risk theory. *ASTIN Billetin*, Vol.32 ,No.1, 145-161.

[Mohammad, 2000] Mohammad A. Al-Fawzan (2000) . Methods for Estimating the Parameters of the Weibull Distribution . *King Abdulaziz City for Science and Technology* mfawzan@kacst.edu.sa.

[Mohamed et. al, 2010] Mohamed Amraja Mohamed, Ahmad Mahir Razali and Noriszura Ismail (2010). Approximation of Aggregate Losses Using Simulation . *Journal of Mathematics and Statistics* , Vol. 6 ,No. 3, 233-239.

[Neuts, 1981] Neuts, M.F. (1981). Matrix-geometric solutions in stochastic models. *Series in the Mathematical Sciences. Johns Hopkins University Press* Baltimore, Md hNps://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=2737.

[O'cinneide, 1990] O'cinneide C. A. (1990). Characteristics of Phase-type distribution. *Comm Statist Structure model*, No.6,1-57

[Oketch, 2011] Oketch N. (2011) . Negative Binomial distributions for fixed and random parameters . *Nairobi repository* University of Nairobi.

[Oscar, 1981] Oscar Garcia (1981) . Simplified method of moments estimation for the Weibull distribution . *New zealand Journal of Forestry science* Vol. 11, No. 3 pp. 304-306.

[Paul et.al, 2010] Paul Embrecht and Marco Frei (2010). Panjer Recursion Versus FFT for Compound Distributions. *Seminar For Statistics*, pp.1-13.

[Pavel, 2011] Pavel V. Shevchenko (2011). Chapter 3 calculation of Compound Distribution. *Springer Science and Business Media LLC*.

[Pavel, 2010] Pavel .v. Shevchenko(2010).Calculation of Aggregate Loss Distribution.*The Journal of Operational Risk*,5(2), pp. 3-40.

[Persi et. al, 1991] Persi Diaconis and Sandy Zabell (1991). Closed Form Summation for Classical Distributions: Variations on Theme of De Moivre Networks. *Journal of Statistical Science* , Vol 6, No. 3, 284-302.

[Rama et. al, 2016] Rama Shanker and A. Mishra (2016) . A Two-Parameter Poisson-Lindley Distribution . *International Journal of Statistics and Systems* Vol. 9, No. 1, pp. 79-85.

[Rama et. al, 2015] Rama Shanker and Hagos Fesshaye (2015). On poisson-lindley distribution and its applications to biological sciences . *Biometrics and Biostatistics International Journal* Vol. 2, No. 4, pp. 103-107.

[Rama et. al, 2017] Rama Shanker and Kamlesh KUmar Shukla (2017). A Zero-Truncated Two-Parameter Poisson-Lindley Distribution with an Application to Biological Science.*Turkiye Klinikleri J Biostat*,9(2)85-95,hNps://www.researchgate.net/publication/319861343.

[Robertson, 1992] Robertson.J.(1992). The Computation of Aggregate Loss Distribution. *In Proceedings of the Casualty Actuarial Society*,Vol. 79,No. 150, pp.57-133.

[Antony et. al, 2020] Rono A. , Ogutu C. and Weke P. (2020). On Compound Distributions for Natural Disaster Modelling in Kenya.*International Journal of Mathematics and Mathematical Sciences*, vol. 2020, Article ID 9398309, 8 pages, 2020. hNps://doi.org/10.1155/2020/9398309.

[Serkan, 2016] Serkan Eryilmaz (2016) . Compound Markov negative binomial distribution. *Journal of Computational and Applied Mathematics* Vol. 292, pp. 1-6.

[Sietske, 2009] Sietske Greeuw (2009) . On the relation between matrix-geometric and discrete phase-type distribution . *Amsterdam* University of Amsterdam

[Sophie et. al, 2012] Sophie Hautphenne, Guy Latouchey and Giang T. Nguyen (2012). On the nature of Phase-Type Poisson distributions .*Annals of Actuarial Science*,Vol. 8,No. 01, pp. 1-25.

[Stuart et. al, 2008] Stuart A. Klugman, Har ry H. Panjer, Gordon E. Willmot (2008). Loss Models.*Wiley*, hNps://www.wiley.com/en-gb/Loss+Models 3A+From+Data+to+Decisions 2C+5th+Edition-p-9781119523789 .

[Sundberg, 1974] Sundberg (1974). Maximum likelihood theory for incomplete data from an exponential family.*Scandinavian Journal of statistics*, Vol 1,No. 2,49-58.

[Sundt et. al, 1981] Sundt, Bjorn, Jewell, Willam S. (1981). Further results on recursive evaluation of compound distributions. *ASTIN Bulletin* Vol. 12, No.1, 27-39.

[Tanka, 2016] Tanka Raj Adhikari (2016). Theoretical Description of A Reparamatrization of Discrete Two Parameter Poisson Lindley Distribution for Modeling Waiting and Survival Times Data. *Tribhuvan university Journal*, Vol. 19, No. 1, pp. 217-222.

[Verbelen, 2013] Verbelen R. (2013). Phase-type distributions and mixtures of Erlangs. A study of theoretical concepts, calibration techniques and actuarial applications.*KU Leuven*,Vol. 1,pp. 1-132.

[Vilar et. al, 2009] Vilar J.,R. Cao, M.C Ausin and C. Gonzalez-Fragueiro (2009). Non parametric analysis of aggregate loss models.*Journal of Applied Statistics* Vol. 36.

[Xueyuan, 2010] Xueyuan Wu and Shuanming Li (2010). Matrix-Form Recursion for a Family of Compound Distribution. *Astin Bulletin*, Vol. 40, No. 1, pp. 351-368.

[Younes, 2012] Younes Mouatassim (2012). Zero-modified discrete distribution for operational risk modeling.*The Journal of Risk Finance*,Vol. 13,No. 5,pp. 476-490.

Appendices

Appendix I

This appendix shows MATLAB codes used in estimating matrix parameters of phase type distributions.

MATLAB codes for matrix parameters

```
%THREE STATE

a=850;
b=103;
c=(a-b)/a;
d=62;
e=(b-d)/b;
A=[c,1-c,0;0,e,1-e;0,0,1];
B=A*A*A*A*A*A*A*A*A*A*A

% FOUR STATE

 %Four State
f=850;
g=179;
h=(f-g)/f;
j=127;
k=(g-j)/g;
l=13;
m=(j-l)/j;
B=[h,1-h,0,0;0,k,1-k,0;0,0,m,1-m;0,0,0,1]

H=B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*
B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B*B
```

```
%FIVE STATE

n=850;
p=139;
q=(n-p)/n;
r=85;
s=(p-r)/p;
t=28;
u=(r-t)/r;
v=12;
w=(t-v)/t;
C=[q,1-q,0,0,0;0,s,1-s,0,0;0,0,u,1-u,0;0,0,0,w,1-w;0,0,0,0,1]
M=C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C*C


%SIX STATE
x=850;
y=438;
z=(x-y)/x;
E=384;
F=(y-E)/y;
G=325;
H=(E-G)/E;
J=182;
K=(G-J)/G;
L= 24 ;
M= (J-L)/J ;
D=[z,1-z,0,0,0,0;0,F,1-F,0,0,0;0,0,H,1-H,0,0;0,0,0,K,1-K,0;0,0,0,0,M,1-M;0,0,0,0,0,1]
Q=D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D
*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D*D;
```

Appendix II

This appendix shows MATLAB codes used in calculating frequency probabilities of phase type distributions.

MATLAB codes for frequency probabilities

```
% Phase type Poisson distribution

% 3 state model

A=[0.8788,0.1212,0;0,0.3981,0.6019;0,0,1.0000];
B=[1,0,0;0,1,0;0,0,1];
C=-1*A;
D=expm(C);
E=A.^5;
F=D*E;
H=factorial(5);
I=[F./H]
J=[0.2122,0.0535,0.7343];
K=[1;1;1];
L=[J*I*K]

% Estimation of moments of phase type Poisson distribution

%  3 state model

% Mean

A=[0.8788,0.1212,0;0,0.3981,0.6019;0,0,1.0000];
J=[0.2122,0.0535,0.7343];
K=[1;1;1];
L=[J*A*K]

% Variance

A=[0.8788,0.1212,0;0,0.3981,0.6019;0,0,1.0000];
J=[0.2122,0.0535,0.7343];
K=[1;1;1];
L=[J*A*K]
```

```
% Phase type Negative Binomial distribution

 % 3 state model

% Calculation of alpha

A=[1,0,0;0,1,0;0,0,1];
B=[0.8788, 0.1212,0;0, 0.3981, 0.6019; 0,0,1];
C=[A-B];
E=[7.2169,-1.0000,0;0,0.6496,-1.0000;0,0,0];
F=[0.2122,0.0535,0.7343];
G=[1;1;1];
H=[F*E*G]
I=111/H

  alpha=76

% Frequency probabilities

A=5+76-1;
B=5;
C=nchoosek(A,B);
D=[1,0,0;0,1,0;0,0,1];
E=[0.8788,0.1212,0;0,0.3981,0.6019;0,0,1.0000];
F=[E.^5];
J=[0.2122,0.0535,0.7343];
K=[1;1;1];
L=J*I*K;
M=L*C
```

Appendix III

This appendix shows MATLAB codes used in calculating severity probabilities.

MATLAB codes for severity probabilities

```
% Severity probabilities

% Weibull

% 3 state

A=[1375000/150507.7436];
B=[A*0.47*-1];
C=[2.718281828^B]
D=[1-C]

% Generalized Pareto

% Determining d

d=0.268;
A=[3*d];
B=[1+A];
E=[2*d];
G=[F^0.5];
H=[D*G];
I=[H/B]

% Determining e

d=0.268;
A=[1+d];
B=[2*d];
D=[C^0.5];
E=[A*D];
F=[3118753.791*E]
```

```
% Probabilities

% 3 state

A=[1375000-339721.8];
B=[17.59*A];
C=[B/107440000];
E=[1/17.59];
F=[D^E]
G=[1-F]

% Pareto

% Estimation of alpha

% 3 state

a=2.285714286;
A=[a-1];
B=[2*A];
C=[a];
D=[B/C]

% Probability calculation

% 3 state

A=[436786.4286 + 1375000 ];
B=[436786.4286];
C=[B];
D=[C^2.285714286];
E=[1-D]
```

List of Publications

The manuscripts prepared, submiNed, accepted and published are:

1. Cynthia Mwende, Patrick G. Weke, Davis N. Bundi and Joseph M. Otieno (2021). Estimation of aggregate losses of secondary cancer cases using PH Panjer class (a,b,1) distributions, Afrika Statistika, Vol. 16, No. 4, pp. 3041-3059. DOI: 10.16929/as/2021.3041.194
2. Cynthia Mwende, Patrick G. Weke, Davis N. Bundi and Joseph M. Otieno (2021). Aggregate Loss Distribution for Modeling Reserves in Insurance and Banking Sectors in Kenya, Far East Journal of Theoretical Statistics, Vol. 62, No. 1, pp. 17-34. hNp://dx.doi.org/10.17654/TS062010017
3. Cynthia Mwende, Patrick G. Weke, Davis N. Bundi and Joseph M. Otieno . Phase Type Zero Truncated Poisson Lindley Distributions and their application in modeling Secondary Cancer Cases, Afrika Statistika, accepted and awaiting publication.
4. Cynthia Mwende, Patrick G. Weke, Davis N. Bundi and Joseph M. Otieno . Matrix Recursions for Panjer class (a,b,0) and Its Application in Modeling Aggregate Claim Losses of Secondary Cancer Cases, Advances and Applications in Statistics, accepted and awaiting publication.