UNIVERSITY OF GHANA

AN APPLICATION OF MARKOV CHAINS AND MIXED POISSON DISTRIBUTION IN MODELLING NO-CLAIM DISCOUNT SYSTEMS FOR MOTOR INSURANCE DATA

BY

AKUMINGE NSOH CLEMENT

(10704125)

A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF PHILOSOPHY TO

UNIVERSITY OF GHANA

MPHIL ACTURIAL SCIENCE

MAY, 2022.

DECLARATION

This thesis is my original work and has not been submitted for a degree award in any other university or academic institution.

.... 03/10/22

Akuminge Nsoh Clement Date

This thesis has been submitted for presentation with my approval as the university supervisor.

.... 03/10/22

Dr. Lotsi Anani Date

.... 03/10/2022

Dr. Richard Minkah Date

ABSTRACT

Most No-Claim Discount (NCD) systems are unfair to either one or both parties. Most systems are the simple random walk model, whereby in case of a claim, the policyholder moves down a discount level and vice versa. Modelling of data of claim amount is of paramount importance to manage risk reserve for payment of claims. Actuaries model uncertainty using probability distributions. The movements within the NCD systems are those of the in-between cases, and a fair NCD system, should take into consideration the frequency of claims and the non-homogeneity factor. In this study, the Markov chains have been employed to explain the movement between levels in the NCD system and mixed Poisson distribution to calculate the probabilities, with the mixing distributions been the exponential and the gamma, and Poisson models. The motor insurance claim data from Sweden was used in this study. The study found that the Geometric distribution model was better fitted to the observed claim frequencies in both the maximum likelihood estimation and method moments than the Poisson model. The results for the 3-Level NCD systems showed that the policyholders were rewarded with approximately 95% chances of moving to the next higher level towards attaining a no premium zone if a claim is not made in the cycle and were punished if a claim is made by dropping to the lower level with approximate probability of 0.048 and it was generally observed. The study found that due to the fairness of the system, policyholders who make claims are equally punished by dropping from their current level to the lower one or are made to stay in their current level in the next cycle. In conclusion, the multi-level NCD system was designed to reward policyholders who are extra vigilant on the road in terms of avoiding road accidents and penalized bad road users. The descriptive statistics of the claim frequency revealed that the number of claims by policyholders was right or positively skewed with a mean number of claims approximately zero (0.053).

DEDICATION

To almighty God who is the source of my strength and wisdom, my father and brothers and my boss Engineer Nimako and all my lecturers of this course and my friends and loved ones.

ACKNOWLEDGEMENT

To the Almighty God, who has guided me all through, I thank You Lord for bringing me this far and guiding me through my studies. My Dear Lord, the journey indeed was very tough but You made it possible for me and I am very much grateful to You, Lord.

My special gratitude also goes to my supervisor, Dr. Lotsi Anani, for his support, experience and encouragement that led to the successful completion of this thesis work.

I am also very grateful and thankful to Dr. R Minkah, my co-supervisor and Dr. Louis Asiedu, Dr. Gabriel Kallah-Dagadu and Prof K. Doku-Amponsah for their brotherly love, support, encouragement and advice they offered me during the period of my thesis work.

I would like to also show my appreciation to all my lecturers as well, to whom I am very grateful for their assistance and encouragement and love they showed me throughout my studies.

Finally, I want to appreciate my family and friends and course mates for the support and encouragement they gave me throughout this period.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

The no-claims discount (NCD) system is a means of rewarding insured motorists who have not made a claim within a period for which premiums were paid, by reducing the premiums, or awarding discounts on premiums payable in the subsequent insurance period. In this system, the movement of insured motorists across discount levels is governed by the frequency of the claims; it ignores the severity information, this implies that the future loss of insureds can be modeled with the claims frequency alone, (Kiprotich et al., 2020).

This reduction is a much sought-after discount by all motorists. Just as discounts are a common incentive in this system, penalties are equally popular. Motorists who made claims within a period of insurance for which premiums were paid, would be liable for an increase in their next payment of premiums, (Azaare & Wu, 2020). This incurred increment, known as loading in the context of auto insurance, could be a bane in the pockets of motor insurance policyholders. In effect, the NCD system as described by Anyango (2015), is a multi-layer premium policy which assigns a variable premium class to a policyholder depending on their most recent claim history.

Ideally, it is suggested by Ogeno (2015) that variables such as age; occupation; gender; number of years in driving; number of claims per policy (past claims history); type and use of car and even the place of garage, be used in classifying various automobile risks into homogeneous groups. This is done by a method known as priori rating. Ogeno (2015) goes further to state that NCD systems are either set up by free markets or imposed by state institutions, which more often, than not, treat all policyholders as a homogeneous group, thereby disregarding their individual characteristics.

Additionally, drivers at the time of a road traffic accident may not agree on who is at fault of causing the collision; it is the duty of the insurers of both drivers to find out the driver-at-fault of the claim. The Competition Commission, the statutory body responsible for investigating regulated industries under competition law in the United Kingdom cites the establishment of fault, at the time of a collision between two driving parties, as a factor in determining which driver is penalised by an increment in payable premiums. In situations where fault is not properly established, or where fault is failed to be established, penalties may be wrongly or unfairly meted out to policyholders. Such instances have often led drivers to not only doubt the effectiveness of NCD systems, but refrain from purchasing motor insurance, entirely (Competition Commission, 2012). In Portugal, policyholders terminate their motor insurance policies, once claims have been paid to them by their motor insurer, and move on to purchase new motor polices from a competitor of their former insurers (Gueirrero & Mexia, 2002). This behaviour is not exclusive of policyholders in Portugal, but pertains to policyholders elsewhere, including Ghana.

1.2 Problem Statement

Models are usually employed to decide when a claim be made and how much be paid (Burney et al. 2022). Therefore, modeling of claim amount is an important technique for actuaries to estimate parameters of the data foe the proposed model and making decision for loses and premium calculations (Rafal, 2010). Though an ideal NCD system should be efficient and competitive, seeking to relate an insured's driving experience to the premium chargeable, however, this is far from the case in the real context of most countries, including Ghana. As identified by Azaare and Wu (2020), the No-Claims Discount systems used in Ghana fails to acknowledge both the frequency and severity of policyholders' claims in their designs. There have been extreme cases where policyholders who made claims had to lose all their levels of discount, earned over their incredible years of driving experience, and begin from the no-discount premium level. The treatment of all policyholders in an insurer's portfolio as a homogeneous group without regarding the individuals' peculiarities for the drivers insured (in the case of motor insurance), usually results in this problem.

Also, there is an ineffective movement of information on policyholders amongst insurance companies in the case that a policyholder makes a claim and is liable for a premium load in their subsequent insurance coverage. This implies that, presently, the model allows a policyholder, after claiming during their insurance coverage with one insurer, to move to different insurer (competitor) for renewal, thus enabling avoiding the premium increment. Due to these challenges with the current system, this study examines the desirability of this multi-layer premium system (NCD system). Nath and Sinha (2014) argue that the basic framework should consider a discrete time parameter Markov chain, where the state space should be of different levels of the premium, and the state of a particular insured shifts randomly a year to the next. The randomness of the transitions would follow a probability law of causing accidents each year, hence the percentage of good drivers can be obtained to receive the fully discounted rate in the long run (Nath & Sinha, 2014). The European Federation of Road Traffic Victims (FEVR) is an association of road victim support and advocacy groups from 11 countries with associate numbers from low income countries including Argentina, Ghana etc. In 1997, FEVR produced Impact of Road Death and Injury, a report which included surveys from over 1300 victims, collected by sixteen organizations in nine European countries. Bereaved families and the injured were asked about their level of satisfaction with the various authorities including coroners, police and insurers.

In general, respondents were unsatisfied with the damages offered, the medical examination process, lack of financial justices and the time involved in civil proceedings. Road traffic victims in Switzerland reported the highest rate of satisfaction with insurance companies but even there 60 percent were dissatisfied. UK victims reported considerable dissatisfaction with insurance companies. Overall, 86 percent of bereaved families and 90 percent of the families with a member disabled from a crash expressed dissatisfaction with the UK insurance companies and their reasons included delays in payments of compensation and inadequate amounts involved. Lack of trust in the compensation process will encourage abuse of the system. The insurance companies in Ghana have complained of inflated disability claims reported by doctors. From the insurer's side, both Ghana and South Africa insurers have expressed concern over the liability of having unlimited benefits for road casualties. There is also a problem with the lump sum payment as these are often spent within several years and the victim returns to the state for benefits. The FEVR study estimated the average time taken to settle the claim of a road fatality at 2.6 years and even longer for claims to be paid involving the disabled and these delays will definitely put burden on the affected families financially.

1.3 Objectives of the Study

This study aims at modeling a multi-level discount system that analyses the problem of homogeneity in the treatment of motor insurance policyholders in a typical No-Claims Discount system.

1.3.2 Specific Objectives

1. To analyse the various levels of discount in 3, 4 and 5 premium NCD systems using Markov chains.
2. To model frequency distribution of motor insurance claims using the mixed Poisson distribution and compute the respective transition probabilities of policyholders in the corresponding 3,4 and 5 discount levels NCD systems.
3. To evaluate the hypothetical cases considered in the 3,4 and 5 discount levels for NCD systems.

1.4 Brief Methodology

This study, will employ Markov chains techniques to determine the transition states of insured drivers (policyholders) from one premium discount level to the other. Mixed Poisson distributions involving one-and-two-parameter gamma and exponential distributions will also be used in describing the claims frequency within a period of insurance coverage.

In this work, the *method of moments* as well as the *maximum likelihood estimation* (*MLE*) method would be used to estimate the population parameters of interest, that is the expected values. The data on the Motor Insurance Claims in Sweden, used in this study are predominantly secondary, cross-sectional and describes third party automobile claims for the year 1977. The data lists the number of claims (frequency) and the sum of payments (severity) in Swedish kroners based on 5 categories of distance driven by vehicle, which have further been divided into 7 geographic zones, 7 subgroups of driver claims history and 9 groups of automobiles. They were obtained from an open source: Kaggle, a data science company and a subsidiary of Google.

1.5 Relevance of the Study

This study is a representative illustration of the applications of Markov chains and mixed Poisson distribution in the insurance and finance sectors. It serves as a model for auto insurers, and other risk enterprises that deal with premium discounts, in running a No-Claims Discount that is fair to insurers and their policyholders as well as competitors.

Also, it ensures a hunger for bonus (discount) among drivers that would spur on insureds to make fewer claims thereby encouraging safe driving on our roads. This will be the surest bet on reducing the cost of motor insurance.

1.6 Limitations of the Study

Though the study intends for the best, it is still liable to flaws and constraints; most of which have got to do with implementation of and compliance to insurance regulations. Below are the major constraints worth giving attention:

Policyholders are regarded by insurers to be a homogeneous group; though policyholders are a diverse group of drivers belonging to different risk classes depending on their driving experience, claims experience and age bracket, or even gender.

Motor insurers in their quest to dominate the auto insurance market offer unreasonable discounts to new clients signing onto their policies even to the point of disregarding NIC and GIA regulations on discount limits.

Data on auto insurance claims frequency on Ghanaian insured motorists are not readily available and in cases when they are, they are not readily accessible. Whether this is as a result of motor insurance being unpopular among Ghanaian motorists or due to the NIC and GIA's inability to retrieve data from auto insurers, it is time we compiled an accessible insurance database.

Justification for the Use of the Foreign Data.

It is always a great swish and desire for every researcher to use the data in the country that he is doing the research work, however there are certain compelling reasons why the researcher uses another country's data to do his analysis in his country. My thesis is not an exception, this is, because the data on auto insurance claims frequency on Ghanaian insured motorists are not readily available and in cases when they are, they are not readily accessible. Whether this is as a result of motor insurance being unpopular among Ghanaian motorists or due to the NIC and GIA's inability to retrieve data from auto insurers. Not being able to access data on auto insurance in Ghana and I therefore decided to search for data on the topic and I got access to data available in Sweden and very good data suitable for my research work analysis. Therefore data used for the analysis was readily available and readily accessible in Sweden.

1.7 Organisation of the Study

The outline of this study, on the application of Markov chains and mixed Poisson distribution on No-Claims Discount systems in Ghana, is given below.

This section provides brief descriptions on the Literature Review, Methodology; Results and Discussions and Conclusion and Recommendations.

A review of past and contemporary works, as well as brief explanations of keywords relevant to this study, can be found in the Literature Review.

The Methodology involves a detailed look at the statistical methods, tools and motor insurance data used to solve the problem of homogeneity mentioned in section 1.3.1 in the form of multiple case studies. Here, hypothetical NCD systems are described under 3,4 and 5 discount-levels. The Maximum Likelihood Estimation (MLE) and Method of Moments are the main statistical methods used in estimating the parameters of the Poisson and mixed Poisson-exponential distributions. The resulting estimators and probability functions from the Methodology are then used, together with the descriptive statistics of the secondary data on Swedish Motor Insurance claims, to compute the transition probabilities for the Markov chain models in the multiple NCD systems considered in the case study. The final transition matrices for all hypothetical cases, under the 3,4 and 5 discount-level NCD systems, are then assessed and used to determine which systems are strict, fair and ideal. Finally, the Conclusion and Recommendations gives a concise summary of the entire work and the salient findings, as well as limitations, identified in the course of the work that need attention by scholars considering this field of study.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

The rules applied in a particular NCD system depend on the regulations. There is the free market, where there is total freedom and insurance companies design their own NCD systems using their own rules. Then there are the government imposed systems, where the rules that are applied in the NCD systems are governed by government and every insurer has to apply these rules. What follows is the literature review on different rules applied in both types of systems used by different countries.

This part of the study consists of a body of existing works undertaken by both past and contemporary scholars on the NCD systems. It shows the links and contrasts in the methods, statistical tools and objectives in the various works and the most favourable methods are determined.

Works on the general applications of Markov chains and mixed Poisson distributions related to this study are briefly discussed. It also highlights on the keywords that are mentioned in core sections of the study that are worth defining.

2.2 Overview of Motor Insurance in Ghana.

In Ghana, the National Insurance Commission (NIC) is the major body responsible for regulating and supervising all insurers in the industry as stipulated by the Insurance Act, 2006 (Act 724). The NIC, in accordance with the Act, has the legal right to enlist all insurance companies, together with their products and intermediaries, fit for operation in the Ghanaian insurance market. Thus, these are the only motor insurance policies permitted by the NIC in Ghana:

1) Third Party Liability Policy

2) Third Party Fire and Theft Policy

3) Comprehensive Motor Insurance Policy.

However, the two main motor insurance policies highly purchased by policyholders on the Ghanaian insurance market are the Third Party Liability and the Comprehensive Motor Insurance Policies.

Third Party Liability Policy, as stated by Gönülal (2010) uses actuarial models to ensure that damage to third party health and property caused by an accident for which driver and/or owner of the car were responsible is covered.' A policy may be taken out by the owner of a vehicle or by a lawful possessor authorized by the owner on behalf of the owner.

Compulsory Insurance, under the Third Party Liability Policy, is a financial protection system built to prevent any grievance that third parties could face, due to lack of solvency of

first party who caused bodily injury or property damage following any event related to a car accident. What Gönülal (2010) failed to look at in his study is the First and Second Party Policyholders who form the majority of the population. His study did not also look at the number of no at-fault claims. According to him the number of at-fault claims is assumed to follow Poisson distribution. This study will look at both First and Second Party Policyholders and the number of no at-fault claims and the distributions they follow.

On the contrary, a Comprehensive Policy provides the benefits available under the third party insurance cover and in addition, it protects the insured's own vehicle against the risk of fire (explosion and lightning), flood, earthquake, malicious damage, theft or attempted theft and damage to the insured's vehicle in use by collision and overturning (Kumaga, 2016).

Kumaga (2016) further mentions that insurance companies focus more on selling comprehensive motor insurance than third party policies because it generates higher revenue.

Lemaine and Zi (1940}, compare the merit rating systems of different countries in the third party automobile insurance rating. They simulate and compare systems of different countries with different stages, levels and rules, using Stationary Average Premium level, the variability of the policyholder payments, their elasticity with respect to the claim frequency and the magnitude of the hunger for bonus. The number of at-fault claims for a particular driver is assumed to conform to a Poisson distribution. Their study failed to look at what happens to no at-fault claims for a particular driver and what distribution is assumed to conform to the number of no at-fault claims for a particular driver. This study will look at number of no at-fault claims and the distribution that is assumed to follow it.

Emmamverdi et al (2013), compare the system being conducted by the Iranian insurance companies. The system is a 5 by 5 stage NCD system and is assumed to follow the Poisson distribution. They present a technique for coming up with optimal scales in automobile insurance that can be commercially implemented and have reasonable penalties. The challenge of their study is that they failed to give suggested penalties that they talk about in their study even though they also use Poisson distribution in their study and their study did look at 3 by 3 stage NCD system and 4 by 4 stage NCD system, therefore this study covers stages 3, 4 and 5 in NCD system.

Nath and Sinha (2014), inspect the desirability of the multi-layer premium system of the Insurance and Regulatory and Development Authority of India, with six levels. They discover that though bad drivers are twice as likely to claim as good drivers, the premium they are charged is only on average marginally higher. They suggested that the levels of the NCD should be adjusted so that bad drivers pay double premium that of good drivers.

2.3 Class-Based Premium Rating in NCD Systems in Ghana

Automobiles under insurance coverage are assigned classes on the basis of their use. These classes then serve as determinants or factors in the primary rating of the insurer's risk analysis. Azaare and Wu (2020) mention that the No-Claims Discount system in Ghana, officially recognises three classes of cars. Among the existing classes, and other emerging ones, are the following: 'driving to work', 'pleasure use', 'farm use', and 'business use'.

'Pleasure use' is so-called if a car is used leisurely or occasionally.

Automobiles assigned to the 'pleasure use' class pay lesser in premiums as compared to vehicles allocated to the class for 'business use'. Since pleasure use cars are driven fewer times in a given period as opposed to business use vehicles which may be driven on a daily; the *claims exposure* is lower for the former than the latter. They applied the Kullback-Leibler divergence, Kolmogorov Smirnov, Anderson-Darling statistical tests and maximum likelihood estimation (MLE) to estimate policyholder's claims. The results suggest that Ghana's auto policyholder's claims are better approximated using lognormal distribution, the industry can adequately evaluate policyholders' claims to minimize potential loss. Additionally, this distribution could enable the market reach decisions on premiums and expected profits theoretically. Their study ignored the benefits of the policyholders whose contributions that the insurance companies rely on to make their profits. They did also consider the levels of policyholders in their study because the models used can not do proper analysis on the stages or levels. This study therefore considers the various levels 3, 4 and 5. Azaare and Wu looked only at the frequency of claims and severity of the claims and not the frequency and severity of the damages to the policyholders. They did not also consider specificities in their study such as at-fault claims or no at-fault claims, they only considered the generality of claims in their study. This study will look at all these gaps in their study and address them.

Insurers are in the business of servicing claims of their policyholders whenever they duly arise. But they can only do this if premiums are paid by the insureds. Premiums are regular payments made by a policyholder, in this case a car owner, to an insurer in exchange for a claim amount whenever a loss arises as stated in the insurance contract. Premiums charged by insurers vary from one car to the other, depending mainly on the use classification.

2.4 NCD Systems: A Chronological Review of Relevant Existing Works.

Walhin and Paris (1999) constructed a bonus-malus model using the parametric method encompassing the mixed Poisson distribution, which they claim is more general than the traditional Negative Binomial or Poisson Inverse Gaussian distributions. They further justify their preference for the parametric approach, to the non-parametric case, by citing its continuity. They however admitted in their conclusion, that the non-parametric method is a better method, in terms of goodness-of-fit, than their three-parameter model.

Prior to Frangos and Vrontos (2001), a chunk of the actuarial literature covering the optimal Bonus-Malus (NCD) system shows that premiums were assigned to each policyholder based on the numbers of accidents recorded in their driving history. Motivated by this unfair both the frequency and severity of claims in calculating the premiums of insured motorists. This helped to cater for the unfairness, of equal premiums, dealt to policyholders who had a relatively small loss resulting from an accident in relation to those who reported huge losses.

Walhin and Paris again, this time in 2001, conducted a more practical study on Bonus-Malus systems using both parametric and non-parametric mixed Poisson distributions. At the end of this improved study, Walhin and Paris (2001) derived an iterative algorithm based on Lemaire's algorithm to compute corresponding premium levels.

Apparently, Lemaire was an actuarial scholar who had come up with an earlier algorithm in 1977. Basing their work on Lemaire (1995) and Baione et al. (2002) identified a progressive reduction of the average premium which causes financial imbalance in the context of the Italian auto insurance market. In that vein, their study is focused on designing an optimal tariff structure that satisfies both transparency and financial balance conditions. They employed the first order Markov chain in this work.

Kliger and Levikson (2002) completed a project on the pricing of no-claims discount systems using in conjunction with Markov decision processes, game theory. Their final paper comes with the novelty of introducing and analysing the interplay between the motor insurer and the insured motorist, providing a pricing tool that considers the policyholder's best response. It is also worthy of notice that Kliger and Levikson consulted Lemaire (1985), Lemaire (1988) and Lemaire (1995) in the construction of their work.

Entirely utilising linear programming in their paper, Heras et. al. (2004) designed Bonus-Malus premium scales with interesting applications in motor insurance. Similar, in objective, to past actuarial scholars such as Baione et al. (2002) and Frangos and Vrontos (2001), they addressed the problem of inequality of premiums payable by policyholders, regardless of severity of claims. Additionally, they look at the monotonicity and proper variability of the premium scale and the improvement of the certain efficiency measure including *RSAL* and the elasticity of the whole NCD system.

The RSAL, or relative stationary average level, is an adequacy measure that measures the severity of the NCD system by expressing the relative premium paid by an arbitrary policyholder as a percentage, when the stationary state has been approached. Like an indispensable force of NCD, Lemaire (1985), Lemaire (1988) and Lemaire (1995) had to be credited in their work.

Pitrebois et al. (2005) observed that the Bonus-Malus System often leads to high 'maluses' any time claims at fault are made. In practice, these maluses are quite difficult, if not impossible, to implement. They therefore illustrate, in their dissertation, how to avoid this drawback by combining a posteriori premiums correction with deductibles depending on a policyholder's level occupied on the premium scale. In the case of claim frequency, the insurance portfolio is assumed to be non-homogeneous whereas homogeneity is assumed for claim severity. Also, the claim frequency is assumed to follow a mixed Poisson distribution and the motor insurance portfolio is closed to incoming and outgoing policyholders.

Detailed micro-level automobile insurance *records* refer to experience at the individual vehicle level, which consists of vehicle and driver characteristics, insurance coverage, and claims experience. Frees and Valdez (2008), in their quest to model the micro-level records of a major Singaporean insurance company, proposed a hierarchical model for three components corresponding to the type, frequency and severity of claims. They came up with the negative binomial distribution for assessing claims frequency. For the type of claim, they suggested the multinomial logit model that is, whether it is third party property damage, third-party injury, insured's own damage or some combination. The third model suggested by Frees and Valdez (2008) is however, the generalised beta distribution.

Applying AHP, Analytic Hierarchy Process, an expert risk scoring method, Chehui et. al. (2011) determine the weights of the various risks factors that influence the incidence of road accidents. They considered risk factors such as the nature of roads, weather conditions, driving time and driver fatigue. They conclude that people, vehicle and environmental factors should be included in the auto insurance underwriting process in China.

However, in Malaysia, the NCD system in force prior to 2013 was claimed to be the simplest system in the world. In all of its simplicity though, Manan et al. (2013) report that the NCD system does not treat policyholders fairly. Good drivers, who are at the highest discount level, are penalised to a larger extent than drivers at the zero percent discount level of the premium scale. In that, if a good driver makes a claim, he/she is reverted to the zero-discount premium level. Manan et al. (2013) therefore proposed for a multiple discount level Markovian NCD system to be implemented to treat policyholders more instead of the abusive one imposed by the Malaysian government.

Questing for the likelihood of claims by different categories of insured motorists, (Nath & Sinha, 2014) divided the policyholders in the Karimganj, North Tripura and West Tripura districts of Thailand into two groups: ‘good drivers’ and ‘bad drivers’. This classification was done purely on a two-year driving experience and accident records of the policyholders. A transition probability matrix was generated for different discount levels to calculate the respective amount of premiums payable in the long term by policyholders in the districts. And not in the least surprising, Nath and Sinha, (2014) was made possible by referring, in no small way, to the works of Lemaire (1985), Lemaire (1988) and Lemaire (1995).

Limiting their work to a three-level-discount premium NCD system, Wu et al. (2015), developed a discrete-time risk model to correlate claims by insured motorists to the payable premiums. For simplicity, only three premium levels were considered. However, recursive formulae are derived to compute the ruin probabilities of policies. In actuarial science literature, ruin probability refers to the likelihood that at a future time the total claims payable by an insurer exceeds the sum of the initial surplus and the total premiums collected up to that time. The correlation between claims and premiums payable therefore allow their impact on ruin probability to be easily determined.

Taking a recap at this point, it needs no say that most of the relevant works on NCD systems are highly based on the pioneering works of Lemaire (1985), Lemaire (1988) and Lemaire (1995). Aside that, these two observations are rather more obvious. In that, from Walhin and Paris (1999) through Kliger and Levikson (2002) to Wu et al. (2015), most of the scholars whose works have been cited are Western or oriental scholars. This shows why more work on NCD systems is needed in an African, and for that matter in the Ghanaian context.

Quite recently in Kenya, Anyango (2015) constructed an NCD system using Markov chains to explain the movements between discount levels of premiums and calculating the probability entries for the transition matrix using mixed Poisson distributions. 3, 4, 5, 6, up to 10 discount levels of premiums were considered in the study, with multiple cases for each premium level. In the second part of the methodology, the Geometric, Negative Binomial, with both one and two parameters estimated by the MLE and method of moments, and the mixed Poisson Lindley distributions were used to fit the claims frequency data. They found that the Geometric and Poisson Lindley distributions compare well with the observed claims frequency data, and therefore recommend that data on current claims experience be used in subsequent works.

To crown it all up, is Azaare and Wu (2020). This study examines that the most current No-Claims Discount (NCD) system used in Ghana’s motor insurance market. It finds the Ghanaian system outdated and ineffective, and therefore puts forward an alternative Bonus-Malus System (BMS) that suits the present Ghanaian motor insurance market. The model constructed in this work acknowledges both the frequency and severity of claims in its design, which checks the problem of good drivers being unfairly penalised for making claims. In all, it is a model worth implementing, and should the NIC, GIA and all insurance stakeholders play their part, motor insurance in Ghana would be much more patronised and beneficial to all.

Institute and Faculty of Actuaries Examination CT4(2009), over the years has several questions with the NCD systems with different rules applied within these systems. From 3 by 3 stage NCD systems to 4 by 4 stage NCD systems, most of which are assumed to follow a Poisson distribution. Most scholars in this literature review use both lognormal and Poisson distribution in their analysis, however the challenge is that most NCD systems are that of the simple random walk model, whereby in case of a claim, the policyholder moves down a discount level and vice versa. Then there are extreme cases, whereby if a driver makes claim(s), he loses all the discounts accumulated over the years and goes to the level of full premium payment so therefore lognormal and Poisson distributions may not give fair calculations on the claim(s). In line with this, this study will employ different techniques and models which includes, Mixed Poisson distributions involving one -and-two parameter gamma and exponential distributions will also be used in describing the claims frequency within a period of insurance coverage.

2.5 Terminologies used in the Study

2.5.1 Markov Chains

A stochastic process is a mathematical model that evolves over time in a probabilistic manner. Samaniego (2003). And probably the most common and highly applied kind of all stochastic processes, the Markov chain is described by statisticians in a number of interesting ways.

Be it the simple explanation that, "a Markov chain is a mathematical system that experiences transitions from one state to another according to certain probabilistic rules. The defining characteristic of a Markov chain is that no matter *how* the process arrived at its present state, the possible future states are fixed."

Or Samaniego (2003) who stipulates that a sequence of trials of an experiment constitute a Markov chain if the outcome of an experiment belongs to a set of discrete states; and the outcome of an experiment depends solely on the present state, and not on the past states.

Or Myers et. al. (2010), who infer, that a Markov chain is a stochastic process whose past and future states are independent when the present is known.

It can therefore be noted from the assertions cited above that all Markov chains have an underlining requirement. This can be stated loosely as: the future is not dependent on the past, but the present. This fundamental condition, which distinguishes Markov chains from all other stochastic processes, is called the Markov property.

2.5.2 The Markov Property.

Deducing from the work of Levin et. al. (2009), the Markov property can be informally explained as the conditional probability of moving from one state, say X, to another, say y, no matter what sequence say, $X_1, X_2, \cdots, X_{t-1}$, of states precede X.

2.5.3 Mixed Probability Distributions.

Mixed type distributions are usually grossly cited with respect to zero-inflated data. Zero-inflated data, as an expression was initiated in the early 1990s mainly in conjunction with discrete parametric distributions such as the zero-inflated Poisson or ZIP distribution, for short. It has now been tied to continuous distributions, (Lambert, 1992) and (Tu & Liu, 2002) Mixed type distributions, having both continuous and discrete components to their probability distribution, are intrinsic components of numerous diverse and vital models. Queuing is a typical case in point. Several meteorological models including instantaneous wind speed supporting turbine placement and daily rainfall supporting hydrological models (NCEI, 2017).

Given the affirmations provided above, it is evident that mixed probability distributions have a countless list of applications across various fields of endeavour. However, this study would exploit the mixed Poisson distribution in conjunction with Markov chains.

2.5.4 Mixed Poisson Distributions

Mixed Poisson distributions have been applied in a large range of science fields to model non-homogeneous populations. They have been broadly used to model observed situations where characteristics reflected by the given data have differed from the anticipated characteristics under a simple component distribution. They can be formed when the Poisson distribution is merged with the exponential, one and two parameter gamma distributions, the Lindley distribution among others. The distribution of the number of car accidents in an automobile insurance portfolio is known to be well fitted by mixed Poisson distribution (Walhin & Paris, 1999).

2.5.5 Claims Experience and Loss Distributions

Claim frequency refers to the total count of claims that occur on a block of policies. The claims experience random variables are often assumed to follow certain probability distributions, referred to as loss distributions (Omari et al., 2018). Loss distributions are typically skewed to the right with relatively high right-tail probabilities.

2.5.6 Severity of Claims

Severity of claims, in motor insurance, can either be the amount paid due to a loss or the size of the loss event. For insurer's to be in a position to pay claims resulting from existing portfolios of policies in future, it is required that the insurer sufficiently models past and current data on claim experience then use the models to forecast the expected future experience in claim amounts (Kiprotich et. al., 2020) Similarly, Omari et al. (2018) define it as the monetary amount of loss on each policy or on the portfolio as a whole.

2.5.7 Premium Rating

Premiums are regular payments made by a policyholder, in this case a car owner, to an insurer in exchange for a claim amount whenever a loss arises as stated in the insurance contract. In the business of insurance, pricing insurance coverage is referred to as premium rating.

2.6 Establishing Fault in NCD Systems

The *Road Traffic Act* is a regulation set to check the movements and general conduct of motorists on our roads and ensure that they meet enforceable standards that should be complied with. The Road Traffic Act, 2004 (Act 683) Ghana and the Road Traffic Act 1988 (RTA, 1988) UK are a few of the many strict regulations that have been enacted by Acts of Parliaments.

The drivers, at the time of a motor collision, may or may not know which driver is at fault. And should they know which party is at fault, the party at fault may not readily admit in the heat of the incident. It is therefore the insurers' obligation to identify which party is at fault and decide on the insurer responsible for the claims that may arise. That said, the drivers are also obliged to report to their insurers or the broker who sold them their policies for due process to follow. The act of immediately notifying the insurer of the accident is known as first notification of loss (FNOL).

In the event of a road traffic accident (RTA), the fault driver is legally entitled to some amount of compensation, which have been stated in the private motor insurance *(PMI)* policy. Should a collision of automobiles occur, the fault driver would need to take on a temporary replacement vehicle (*TRV*) given that affected vehicle may require repair and would be unavailable for some period. Provided a comprehensive policy is in place, he is covered for repairs to his vehicle up to and agreed excess. Otherwise, only the affected third party is compensated.

In the case of a non-fault claim, the non-fault driver is indemnified at the entire expense of the fault driver. This implies that the insurance Law of Indemnity is applied to the letter. The non-fault driver is legally entitled to be reinstated to the financial position he held prior to the collision or accident. And the fault driver (party) is largely responsible for the associated costs. Multiple parties may also be involved in the claims process regarding the non-fault driver's loss. These include: the non-fault broker, the non-fault insurer, the fault insurer, CMC (claims management company), a TRV provider, such as a car hire company and an auto-repairer.

2.7 Applications of Markov Chains Related to the Study of NCD Systems

2.7.1 Measurement of Credit Risk

Markov chains, as shown in a study by Myers et al. (2010), are applied to measure the credit risk of firms and business enterprises. They found that the transition matrix derived from the credit data of firms represent the future evolution of their crediting ratings or scores. This, they claim, will describe the probability that a particular firm, corporation or even country will either remain in a current credit rating or move to a better or worse credit rating.

The main challenge with this application is how to generate the transition matrix from the credit data. The most common way out for most researchers is by estimating credit rates by analysing data from credit rating agencies like Moody's and Fitch and Standard & Poor's 500 (S&P). This method is not very reliable since the future might not develop smoothly as in the past.

2.7.2 Forecasting of Stock and Alternative Asset Trends

Markov chains, in conjunction with their corresponding diagrams (transition graphs), can be used to model the probabilities of stock trends and the patterns of other financial instruments such as treasury bills, government bonds, and crude oil prices.

2.7.3 Customer Relationship Models (CRM)

The customer's lifetime value is a needful concept in the context of interactive marketing. Dwyer (1989), in (Pfeifer & Carraway, 2000), assisted in publicising the lifetime value concept (LTV) by illustrating the computation of LTV for both customer migration and customer retention situations. Here, customer retention refers to situations where customers who no longer patronise a business party are considered as 'lost for good'.

While Dwyer (1989) illustrated his calculations of the LTV model using only two situations of the customer's relationship behaviour, (Berger & Nasr, 1998) show an improved model involving five customer behaviour patterns: four involving customer retention and one involving customer migration. Perhaps, this is a statistical tale of 'two heads are better than one'.

2.7.4 Operations Research in Airline Switching

Operations research, or simply OR, is explained by Hillier and Liebermann, as "research on operations" applied to problems that concern how to conduct and coordinate the activities within an organisation.

In aviation, Markov chains have been used by airlines such as British Airways to analyse switching by their clients. For instance, what are the chances that a current Business Class customer would switch to: Economy Class on their next flight booking; or switch from Economy Class to Premium Economy Class; or even make two or more flight bookings from now; or stay with the Business Class on their next flight; or even switch from one airline to a different airline, altogether?

These probabilities of airline switching help the management of the airline to make decisions concerning what capacity of passengers to expect on a particular class of flight during specified periods. It goes a very long way in estimating their costs, profits and other financial variables.

2.7.5 Brand Loyalty

Mutiu & Dotun (2015), in their study to examine customers' brand loyalty for mobile phones using Markov chains, asserted that, "Markov chains method comes out as the primary and most powerful technique in predicting the market share a product will achieve in the long term especially in an oligopolistic environment and in finding out the brand loyalty for a product. The concepts of marketing studies are thought as discrete from the time and place viewpoint and so finite Markov chains are applicable for this kind of process."

2.7.6 Queuing Theory

A series of linked stations for serving in which each user, after departing from a particular station, can pass into another one or exit from the whole system is called the networked queuing system. The general assumption is that one station cannot at the same time serve two or more arrival entities. If the station is busy, the user has to wait for service. At the very moment when the station becomes free, the entity is taken over from the queuing according to the pre-defined rules (Petrović et al., 2008).

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter presents detailed statistical methodology of the study and comes in two parts. The first part entails three different level-discount premiums with varied conditions under separate situations. Here, the 3-Levels-Discount, 4-Levels-Discount, and 5-Levels-Discount premium scales of the NCD systems are considered with transition graphs and transition matrices for illustration. In the second part, the claims frequency distribution is studied using a mixed Poisson distribution. The Poisson-Exponential distribution is chosen, and its parameter is estimated by the maximum likelihood estimation and method of moments.

3.2 Hypothetical NCD Systems with Multiple Levels of Discount

Three hypothetical cases are considered for each of the three multiple level-discount NCD systems, namely: the 3-Levels, 4-Levels, and 5-Levels NCD systems.

3.2.1 A 3-Levels-Discount NCD Systems

Under this level-discount system, three situations are considered: each involving three levels of discounts.

CASE 1: If a claim is made, one's discount level is reduced to the next lower level of discount or remains at the 0% level. Otherwise, one moves to the next discount level of premium, or stays at the highest discount level. Here the levels of discount are Level 1, Level 2 and Level 3; represented by $\mathbf{L_1}$, $\mathbf{L_2}$ and $\mathbf{L_3}$, respectively.

CASE 2: If a claim is made, one's discount level is reduced to 0%, regardless of the discount level held by one prior to the claim; otherwise, one moves to the next discount level or remains at the highest discount level. Just like levels of discount in CASE 1, the three levels of discount in this case are also represented as $\mathbf{L_1}$, $\mathbf{L_2}$ and $\mathbf{L_3}$, respectively.

CASE 3: Here, the discount level a policyholder enjoys is reliant on both claims and whether the policyholder was at fault in the events that led to those claims. If one makes a claim at fault, he moves to the next lower level or stays at 0%; otherwise, one stays at the same level prior to the claim if one is not at fault. However, one moves to the next discount level if he does not make any claim. If one is at the highest level of discount and does not make any claim each year, one is given a year's worth of premium waiver, only to return to the highest discount level after a year. Unlike CASE 1 and CASE 2, CASE 3 involves a temporary additional level of discount, the No-Premium Level which will be denoted as $\mathbf{N_p}$, whereas Level 1, Level 2 and Level 3 will still be represented by $\mathbf{L_1}$, $\mathbf{L_2}$ and $\mathbf{L_3}$, respectively.

3.2.2 A 4-Levels-Discount NCD Systems

CASE 1: As in the 3-level-discount premium scale, if a claim is made, one's discount level is reduced to the next lower level of discount or remains at the 0% level. Otherwise, one moves to the next discount level of premium, or stays at the highest discount level. Here the levels of discount are Level 1, Level 2, Level 3 and Level 4, represented as $\mathbf{L_1}$, $\mathbf{L_2}$, $\mathbf{L_3}$ and $\mathbf{L_4}$, respectively.

CASE 2: Here, one moves to the next lower discount level if a claim is made but not at-fault; and moves down to the next two available discount levels lower than the level prior to the claim if one is at fault. In the case of no claims however, one moves to the next level of discount; or remains at the highest level. As in CASE 1, the three levels of discount in this case are also represented by $\mathbf{L_1}$, $\mathbf{L_2}$, $\mathbf{L_3}$ and $\mathbf{L_4}$, respectively.

CASE 3: This is an extension of CASE 2. In that, a policyholder is given a premium waiver (100% discount) for a year provided no claim is made when he attains the highest discount level (Level 4). However, the policyholder would return to Level 4 after a year of free motor insurance coverage. Under this case, Levels 1, 2, 3 and 4 are denoted by $\mathbf{L_1}$, $\mathbf{L_2}$, $\mathbf{L_3}$ and $\mathbf{L_4}$, respectively whereas the No-Premium Level is represented as $\mathbf{N_p}$.

3.2.3 A 5-Levels-Discount NCD Systems

CASE 1: As has been the case in CASE 1 of the 4-level-discount case, so is this case. If a claim is made, one's discount level is reduced to the next lower level or remains at the 0% level. Otherwise, one moves to the next level of premium, or stays at the highest discount level. The Levels 1,2,3,4 and 5 of this no-claims discount system are represented by $\mathbf{L_1}, \mathbf{L_2}, \mathbf{L_3}, \mathbf{L_4}$ and $\mathbf{L_5}$.

CASE 2: CASE 2 in the 4-level-discount premium scale applies here, as well, however this system involves five levels of discount instead of four. These levels of discount include Level 1, Level 2, Level 3, Level 4 and Level 5 denoted by $\mathbf{L_1}, \mathbf{L_2}, \mathbf{L_3}, \mathbf{L_4}$ and $\mathbf{L_5}$.

CASE 3: CASE 3 in the 4-level-discount premium scale applies here too. Here, one who attains the highest-level-of-discount receives a premium waiver for one year if no claim is made. The winner of the premium waiver, however, returns to the highest level of premium discount after one year. In addition to $\mathbf{L_1}, \mathbf{L_2}, \mathbf{L_3}, \mathbf{L_4}$ and $\mathbf{L_5}$ denoting Levels 1,2,3,4 and 5, $\mathbf{N_p}$ is used to represent the No-Premium Level in the case.

3.3 Transition Matrices and Graphs for the Multi-Level NCD Systems

The transition matrices, together with their corresponding transition graphs, for the 3-Levels, 4-Levels and 5-Levels are shown in Figure 3.1, Figure 3.2, and Figure 3.3, respectively. T_{ij} is the transition matrices of the i-Level NCD systems in j cases. P_{ij} is the probability of no claim in i-level of NCD systems in j cases and $1 - P_{ij}$ is the probability of claim in i-level NCD systems in j cases.

3.3.1 The 3-Level NCD Systems' Transition Matrices and Graphs

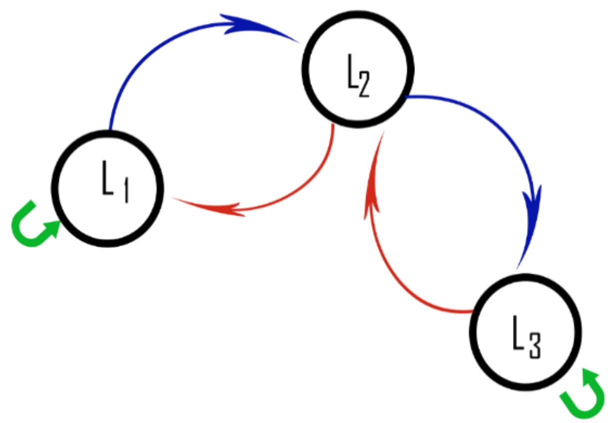
The transition matrices that correspond to the hypothetical NCD systems described in CASE 1, CASE 2 and CASE 3 of the 3-Level-Discount NCD systems in section 3.2 are presented in Equations (1, 2 and 3), respectively.

$$T_{3,1} = \begin{matrix} L_1 \\ L_2 \\ L_3 \end{matrix} \begin{bmatrix} p_{11} & 1-p_{11} & 0 \\ p_{21} & 0 & 1-p_{21} \\ 0 & p_{32} & 1-p_{32} \end{bmatrix} \qquad (1)$$

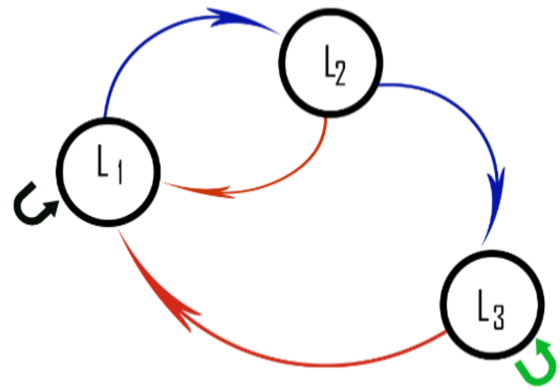
$$T_{3,2} = \begin{matrix} L_1 \\ L_2 \\ L_3 \end{matrix} \begin{bmatrix} p_{11} & 1-p_{11} & 0 \\ p_{21} & 0 & 1-p_{21} \\ p_{31} & 0 & 1-p_{31} \end{bmatrix} \qquad (2)$$

$$T_{3,3} = \begin{matrix} L_1 \\ L_2 \\ L_3 \\ N_p \end{matrix} \begin{bmatrix} p_{11} & 1\text{-}p_{11} & 0 & 0 \\ p_{21} & p_{22} & 1\text{-}\left(p_{21}+p_{22}\right) & 0 \\ 0 & p_{32} & p_{33} & 1\text{-}\left(p_{32}+p_{33}\right) \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad (3)$$

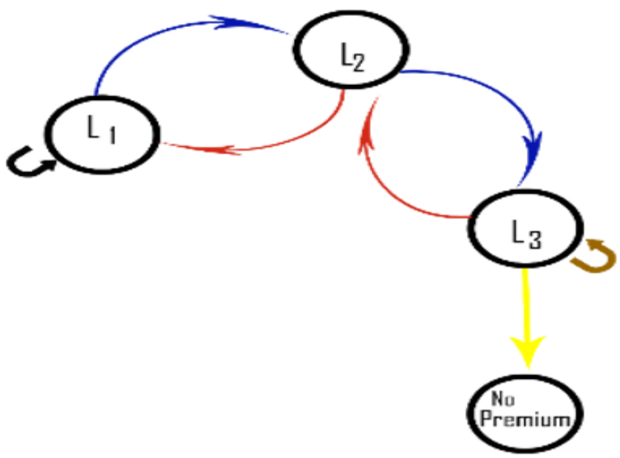
Figure 3.1 presents the transition graphs for the 3-Levels NCD system corresponding to the three transition matrices in Equations (1-3), respectively. CASE 1 denotes the transition graph for Equation (1), CASE 2 denotes the transition graph for Equation (2) and CASE 3 denotes the transition graph for Equation (3).



a. CASE 1



b. CASE 2



c. CASE 3

Figure 3. 1: Transition Matrices Graphs for the 3-Level NCD Systems.

3.3.2 The 4-Level-Discount NCD System's Transition Matrices and Graphs

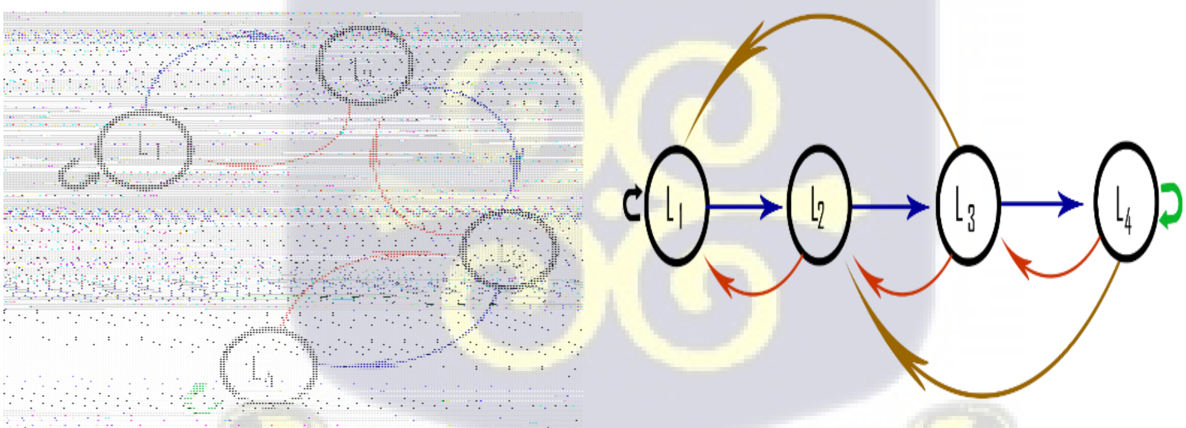
The matrices under the 4-Level-Discount NCD System with the corresponding transition graphs are presented in Equations (4, 5, and 6) and Figure 3.2, respectively. The $T_{4,1}$, $T_{4,2}$ and $T_{4,3}$ represent the transition matrices of CASE 1, CASE 2 and CASE 3, respectively.

$$T_{4,1} = \begin{array}{c} L_1 \\ L_2 \\ L_3 \\ L_4 \end{array} \begin{bmatrix} p_{11} & 1-p_{11} & 0 & 0 \\ p_{21} & 0 & 1-p_{21} & 0 \\ 0 & p_{32} & 0 & 1-p_{32} \\ 0 & 0 & p_{43} & 1-p_{43} \end{bmatrix} \quad (4)$$

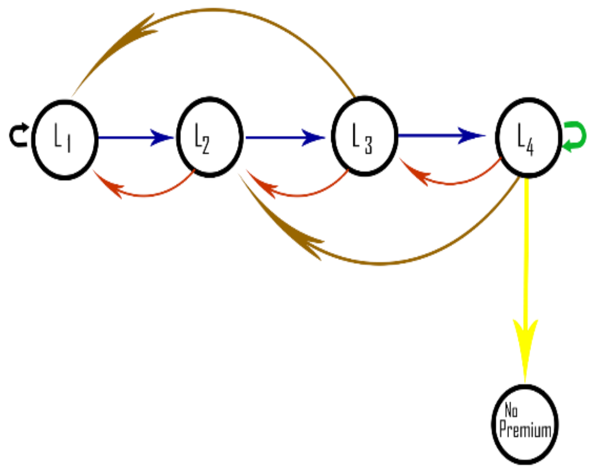
$$T_{4,2} = \begin{array}{c} L_1 \\ L_2 \\ L_3 \\ L_4 \end{array} \begin{bmatrix} p_{11} & 1-p_{11} & 0 & 0 \\ p_{21} & 0 & 1-p_{21} & 0 \\ p_{31} & p_{32} & 0 & 1-(p_{31}+p_{32}) \\ 0 & p_{42} & p_{43} & 1-(p_{42}+p_{43}) \end{bmatrix} \quad (5)$$

$$T_{4,3} = \begin{array}{c} L_1 \\ L_2 \\ L_3 \\ L_4 \\ N_p \end{array} \begin{bmatrix} p_{11} & 1-p_{11} & 0 & 0 & 0 \\ p_{21} & 0 & 1-p_{21} & 0 & 0 \\ p_{31} & p_{32} & 0 & 1-(p_{31}+p_{32}) & 0 \\ 0 & p_{42} & p_{43} & 0 & 1-(p_{42}+p_{43}) \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (6)$$

Figure 3.2 presents the transition graphs for the 4-Level NCD system corresponding to the three transition matrices in Equations (4-6), respectively. CASE 1 denotes the transition graph for Equation (4), CASE 2 denotes the transition graph for Equation (5) and CASE 3 denotes the transition graph for Equation (6)



a. CASE 1 b. CASE 2



c. CASE 3

Figure 3. 2: Transition Matrices and Graphs of the 4-Level NCD Systems

3.3.3 The 5-Level-Discount NCD System's Transition Matrices and Graphs

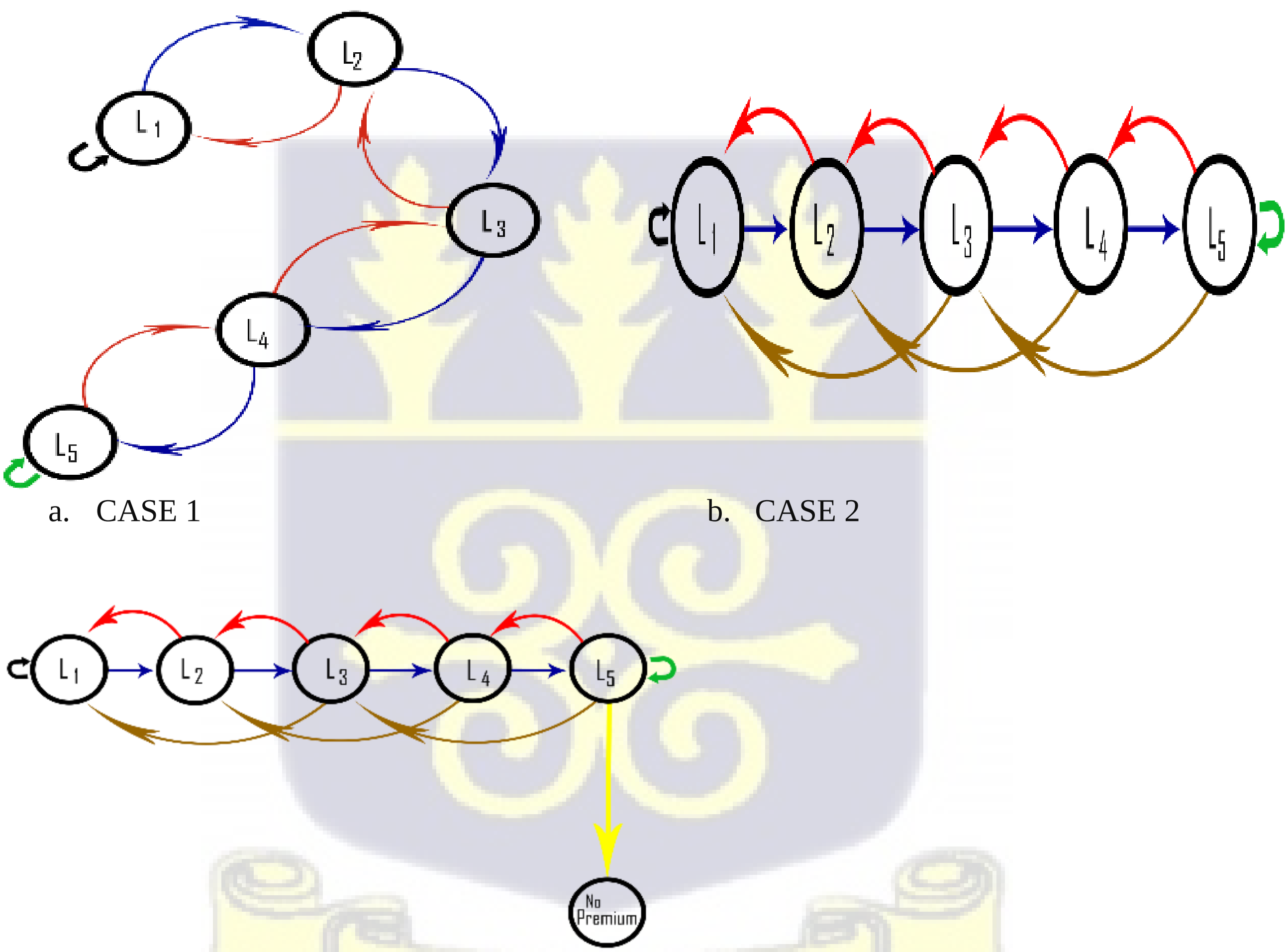
Finally, the 5-Level-Discount NCD Systems transition matrices with their corresponding transition graphs are presented in Equations (6-9) and Figure 3.3. Here, $T_{5,1}, T_{5,2}$ and $T_{5,3}$ denote the transition matrices described in CASE 1, CASE 2 and CASE 3 of the Level-Discount NCD systems.

$$T_{5,1} = \begin{matrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \end{matrix} \begin{bmatrix} p_{11} & 1-p_{11} & 0 & 0 & 0 \\ p_{21} & 0 & 1-p_{21} & 0 & 0 \\ 0 & p_{32} & 0 & 1-p_{32} & 0 \\ 0 & 0 & p_{43} & 0 & 1-p_{43} \\ 0 & 0 & 0 & p_{54} & 1-p_{54} \end{bmatrix} \quad (7)$$

$$T_{5,2} = \begin{matrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \end{matrix} \begin{bmatrix} p_{11} & 1-p_{11} & 0 & 0 & 0 \\ p_{21} & 0 & 1-p_{21} & 0 & 0 \\ p_{31} & p_{32} & 0 & 1-(p_{31}+p_{32}) & 0 \\ 0 & p_{42} & p_{43} & 0 & 1-(p_{42}+p_{43}) \\ 0 & 0 & p_{53} & p_{54} & 1-(p_{53}+p_{54}) \end{bmatrix} \quad (8)$$

$$T_{5,3}=\begin{array}{c} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ N_p \end{array}\begin{bmatrix} p_{11} & 1-p_{11} & 0 & 0 & 0 & 0 \\ p_{21} & 0 & 1-p_{21} & 0 & 0 & 0 \\ p_{31} & p_{32} & 0 & 1-\left(p_{31}+p_{32}\right) & 0 & 0 \\ 0 & p_{42} & p_{43} & 0 & 1-\left(p_{42}+p_{43}\right) & 0 \\ 0 & 0 & p_{53} & p_{54} & 0 & 1-\left(p_{53}+p_{54}\right) \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad (9)$$

Figure 3.3 presents the transition graphs for the 5-Level NCD systems corresponding to the three transition matrices in Equations (7-9), respectively. CASE 1 denotes the transition graph for Equation (7), CASE 2 denotes the transition graph for Equation (8) and CASE 3 denotes the transition graph for Equation (9).



a. CASE 1

b. CASE 2

c. CASE 3

Figure 3. 3: Transition matrices and graphs for the 5-Level NCD Systems.

3.4 Frequency Distribution of Claims

This section presents the frequency distribution of claim amounts. Studies by Lemaire (1985,1988 & 1995) on No-Claims Discount systems, have involved the use of the Poisson distribution to describe the frequency of claims by insured motorists within a year of insurance coverage. However, the mixed Poisson distribution is used lately since it has been observed to factor in the non-homogeneity of policyholders and the claims occurrence. The mixed Poisson distribution is more realistic than the simple Poisson distribution which treats the policyholders as a homogeneous group. The mixed Poisson distribution consists of the Poisson and exponential distribution. Alternatively, the Poisson distribution could have been mixed with the one- and-two-parameter gamma or the Lindley distributions, as is the case in Anyango (2015).

3.4.1 Mixed Poisson-Exponential Distribution

Suppose that the number of claim amount, denoted by X, follows the Poisson distribution with parameter $\lambda > 0$, where λ is also a random variable that follows the exponential distribution with parameter μ. Then the probability distributions are defined as follows:

$$\mathrm{P}\left(X = x \middle| \lambda\right) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \cdots \tag{10}$$

$$f(\lambda) = \mu e^{-\mu\lambda}, \qquad \lambda > 0 \tag{11}$$

Now, using the Total Probability Rule, we have the following

$$\mathrm{P}(X = x) = \int_{-\infty}^{\infty} \mathrm{P}(X \mid \lambda)\, f(\lambda)\, d\lambda \tag{12}$$

Substituting Equations (10) and (11) into Equation (12) and applying the integration to the function, yield the result

$$\mathrm{P}(X=x)=\frac{\mu}{1+\mu}\left(\frac{1}{1+\mu}\right)^{x},\ \ x=0,1,2,3,\cdots \tag{13}$$

The probability distribution function of Equation (13) is a Geometric distribution with the X number of claims.

3.4.2 The Probability Generating Function of the Claim Frequency

In this subsection, the probability generating function for the claim frequency is derived to enable the moments computation theoretically. By definitions, the probability generating function for the number of claims X is given as

$$G_X(t)=E\left(t^X\right)=\sum_{x=0}^{\infty}t^x\frac{\mu}{1+\mu}\left(\frac{1}{1+\mu}\right)^{x} \tag{14}$$

$$\begin{aligned}G_X(t)&=\frac{\mu}{1+\mu}\sum_{x=0}^{\infty}\left(\frac{t}{1+\mu}\right)^{x}=\left(\frac{\mu}{1+\mu}\right)\left(\frac{1+\mu}{1+\mu-t}\right)\\&=\left(\frac{\mu}{1+\mu-t}\right),\ t<(1+\mu)\end{aligned} \tag{15}$$

Now using Equation (15), we can easily obtain the expectation and variance for number of claims, X as

$$E(X)=G_X^{/}(1)=\frac{dG_X(t)}{dt}\Big|_{t=1}=\frac{1}{\mu} \tag{16}$$

Therefore, the expectation of X is $E(X)=\frac{1}{\mu}$, and the variance is given Equation (17) as

$$Var(X)=G_X^{//}(1)-\left[G_X^{/}(1)\right]^2=\frac{1+\mu}{\mu^2} \tag{17}$$

The results of Equations (16) and (17) confirmed that the number of claims frequency follows geometric distribution with parameter $\frac{\mu}{1+\mu}$, where μ is the parameter of the exponential distribution.

3.4.3 Parameter Estimation of the Geometric and Poisson Distributions

The parameters of the theoretical distributions of the number of claims amount would be estimated based on two methods namely, Method of Moments and Maximum Likelihood Estimation (MLE). This subsection presents the two techniques for driving the parameters of the Geometric and the Poisson distributions.

a. The Method of Moments:

The method of moments involves equating the sample moments with the theoretical moments. Hence to estimate unknown parameter (e.g. Mean), an equation is set up by equating the first non-central moment of the population to the sample mean.

Therefore, we have

$$E(X)=\bar{x}, \tag{18}$$

where $E(X)=\frac{1}{\mu}$, from Equation (16) and $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$.

Hence

$$\hat{\mu}=\frac{1}{\bar{x}}, \tag{19}$$

where $\hat{\mu}$ is the estimated value or parameter from the sample.

The moment generating function of the Poisson distribution is given as

$$M_X(t) = E\left(e^{xt}\right) = \sum_{x=0}^{\infty} e^{xt} \frac{\lambda^x e^{-\lambda}}{x!}, \tag{20}$$

Now summing from zero to infinity using Taylor's series identity, we have

$$M_X(t) = \exp\left\{\lambda\left(e^t - 1\right)\right\}. \tag{21}$$

Suppose there is a random sample of size n, where x_i's ; $i = 1, 2, 3, \cdots, n$, then the

$$M_{\sum x}(t) = E\left(e^{t\sum_{i=1}^{n} x_i}\right) = \exp\left\{n\lambda\left(e^t - 1\right)\right\} \tag{22}$$

By the method of moments, the estimator $\hat{\mu}$ for the parameter μ, is estimated by solving the equation:

$$\exp\left\{t\sum_{i=1}^{n} x_i\right\} = \exp\left\{n\lambda\left(e^t - 1\right)\right\} \tag{23}$$

After some manipulations, we have

$$\hat{\lambda} = \frac{\bar{x}\, t}{e^t - 1}, t \neq 0 \tag{24}$$

b. The Maximum Likelihood Estimation:

The method of moments is a simple method of estimating population parameters as compared to the maximum likelihood estimation. However, it may sometimes result in invalid of infinite estimates. A robust but more complicated alternative to the method of moments is the maximum likelihood estimation method. Under this method, the likelihood (or chance) of obtaining the sample, assuming independence among the sample observations is given by

$$L(\mu) = \prod_{i=1}^{n} \mathrm{P}(X = x_i) = \prod_{i=1}^{n} \left(\frac{\mu}{1+\mu}\right)\left(\frac{1}{1+\mu}\right)^{x_i} \tag{25}$$

$$= \left(\frac{\mu}{1+\mu} \right)^n \left(\frac{1}{1+\mu} \right)^{\sum_{i=1}^{n} x_i} \tag{26}$$

Now, taking log on both sides of Equation (26), we have

$$\ln L(\mu) = n \ln\left(\frac{\mu}{1+\mu} \right) + \sum_{i=1}^{n} x_i \ln\left(\frac{1}{1+\mu} \right). \tag{27}$$

By taking the partial derivative with respect to μ and equating the results to zero, we have

$$\frac{\partial \ln L(\mu)}{\partial \mu} = n\left(\frac{1}{\mu} - \frac{1}{1+\mu} \right) - \frac{\sum_{i=1}^{n} x_i}{1+\mu}, \tag{28}$$

$$n\left(\frac{1}{\mu} - \frac{1}{1+\mu} \right) - \frac{\sum_{i=1}^{n} x_i}{1+\mu} = 0$$

$$\Rightarrow \frac{n}{\mu(1+\mu)} = \frac{\sum_{i=1}^{n} x_i}{1+\mu}$$

$$\therefore \hat{\mu} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}. \tag{29}$$

Therefore, the estimate $\hat{\mu}$ for the geometric distribution produced the same in both cases.

The maximum likelihood estimation method for the Poisson distribution is given as

$$L(\lambda) = \prod_{i=1}^{n} \mathrm{P}(X = x_i) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x!} \tag{30}$$

$$= \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}.$$

Taking the natural logarithm on both sides of Equation (30), yields the results in Equation (31)

$$l(\lambda) = \ln L(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i \ln(\lambda) - \sum_{i=1}^{n} \ln(x_i) \tag{31}$$

Now, taking the derivative on both sides with respect to λ, and equating the results to zero yield the required result as:

$$l^{/}(\lambda) = -n + \frac{\sum_{i=1}^{n} x_i}{\lambda}$$

$$\Rightarrow \qquad \hat{\lambda} = \tfrac{1}{n}\sum_{i=1}^{n} x_i = \overline{x} \tag{32}$$

Therefore, the maximum likelihood estimator of the Poisson distribution is the sample mean $\overline{x}$.

The study further discusses the expected claim frequency for both the Poisson and the Geometric distributions. Based on the above resultant formulae, the expected claim frequencies are fitted to the observed claims frequencies under the MLE and Method of Moments methods, for both the Geometric and Poisson distributions.

Suppose □ denote the expected number of the claim's frequencies under the Poisson and Geometric distributions. Then □ can be estimated as

$$\square = N\square\mathbf{P}(X = x), \tag{33}$$

where N is the sum of all observed claim frequencies.

Therefore, the expected number of claims, □ under the Geometric and Poisson distributions are given in Equations (33 and 34), respectively.

$$\square = N\square\left(\frac{\mu}{1+\mu}\right)\left(\frac{1}{1+\mu}\right)^{x} \tag{33}$$

$$\square = N\square\frac{\lambda^{x}e^{-\lambda}}{x!}. \tag{34}$$

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter presents results and discussions of sampled data from the motor insurance claims database. The chapter begins with descriptive statistics of the data where the frequency distribution of the claim count, the mean, mode, standard deviation, and kurtosis were presented. The population estimates of the motor insurance claims were computed based on the MLE and method of moments with reference to frequency distribution of the claim count. The chapter further fitted the observed frequency claims count data to the Poisson and mixed Poisson-exponential (Geometric) distributions separately and their expected claims distributions compared. The chapter concluded with computation of the transition probabilities for the Markov chain of NCD Systems based on the two probability distributions.

4.2 Preliminary Analysis

This section presents preliminary analysis of the sampled dataset employed. The frequency distribution and the descriptive statistics of the number of claim count are presented. Table 4.1 displays the frequency distribution of the number of claims per a policy. It is observed from Table 4.1 that almost 95 percent of the policies recorded zero claims and a little less than five percent recorded one claim. The policies that recorded more than one claims were approximately less than one percent of the total of the total number of policies considered. This result shows the impact of no claim discount in the behaviour of the policyholders.

Table 4. 1: Frequency distribution of number of claim count per policy

Number of Claims per Policy (x)	Claims Frequency	Percentage
0	643953	94.9772
1	32178	4.7460
2	1784	0.2631
3	82	0.0121
4	7	0.0010
5	2	0.0003
6	1	0.0001
8	1	0.0001
Total	**678008**	**100**

Table 4.2 presents the summary descriptive statistics for the number of claims per policy. It is observed that the average number of claims was 0.053 per a policyholder. The modal number of claims per a policyholder was 0.00 with a kurtosis of 181393.490, which indicates a very large measure of outliers in the number of claims. A large value of kurtosis suggests a possible evident of heavy tails for the distribution of the dataset.

Table 4. 2: Descriptive statistics for the number of claims

Statistics	Values
Mean	0.053
Mode	0.00
Standard deviation	0.015
Kurtosis	181393.49

In addition, a pictorial view of the distribution for the number of claims per policy against the claims frequencies is displayed in Figure 4.1. The bar graph depicts graphically the distribution of the dataset used in the analysis of this study. The graph showed a right-tailed distribution which suggests that the motor insurance claims are positively skewed.

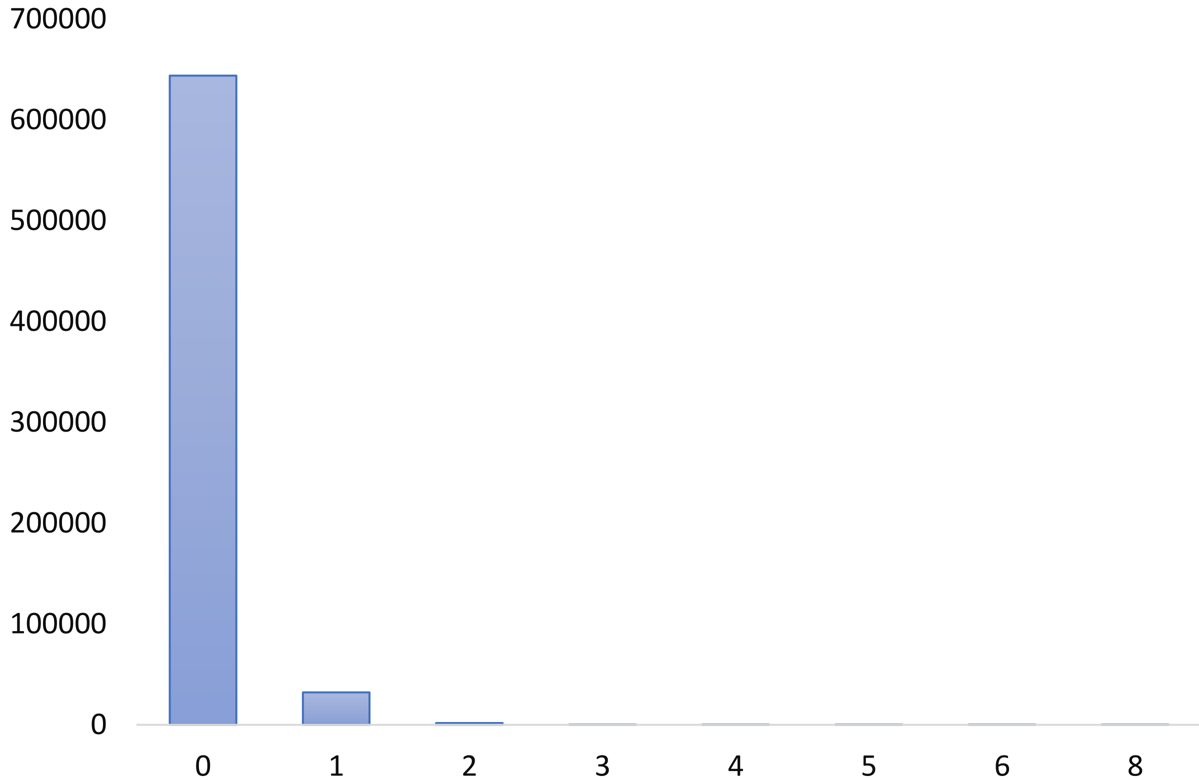


Figure 4. 1: A bar graph showing the distribution of the dataset.

4.3 Fitting Observed Claim Frequencies to Poisson and Geometric Distributions

This section presents number of claims frequencies fit to the Poisson and Geometric distributions using maximum likelihood estimation and the method of moments. Table 4.3 shows the observed number of claims and the expected number of claims for Poisson and Geometric distributions using the maximum likelihood estimation method. It is observed from the Table 4.3 that the expected claim frequencies under the Geometric distribution fit are closer to the observed number of claim frequencies than the Poisson distribution fit. Table 4.4 presents the absolute deviations of the expected claim frequencies of Poisson and Geometric fits under the MLE from the observed number of claim frequencies. It is found that the Geometric distribution fit recorded the smallest absolute mean deviation of 91.312 whilst the Poisson distribution fit recorded a larger value of 517.33. This revealed that the Geometric distribution is best fit to the observed claim frequencies using the maximum likelihood estimation method than the Poisson distribution fit.

Table 4. 3: The number of Claims Frequency distribution under Poisson & Geometric distributions of MLE

No. of Claims	Observed Frequency	Expected Frequency (Poisson)	Expected Frequency (Geometric)
0	643953	642905	643783
1	32178	34178.1	32497.2
2	1784	908.489	1640.41
3	82	16.1	82.805
4	7	0.214	4.18
5	2	0.002	0.211
6	1	2.015	0.011
Total	**678007**	**678008**	**678008**

Table 4. 4: Comparison of absolute deviations of the Poisson & Geometric fit from observed frequencies

Number	Observed Claim Freq.	Expected Number of Claims		Absolute Deviations	
		Poisson	Geometric	Poisson	Geometric
0	643953	642905	643783	1048	170
1	32178	34178.1	32497.2	2000.1	319.2
2	1784	908.489	1640.41	875.511	143.59
3	82	16.1	82.805	65.9	0.805
4	7	0.214	4.18	6.786	2.82
5	2	0.002	0.211	1.998	1.789
6	1	2.015	0.011	1.015	0.989
Total	678007	678008	678008	3999.31	639.193

Mean Deviations	571.33	91.312

Table 4.5 shows the observed number of claims and the expected number of claims for Poisson and Geometric distributions using the method of moments. It is observed that the results of the expected number of claim frequencies based on the method of moments are not different from those observed under the MLE. Thus, the expected claim frequencies under the Geometric distribution fit are closer to the observed number of claim frequencies than the Poisson distribution fit. Furthermore, Table 4.6 shows that the Geometric distribution fit recorded the smallest absolute mean deviation of 91.257 whilst the Poisson distribution fit recorded a larger value of 579.457. This also revealed that the Geometric distribution provides a better fit to the observed claim frequencies based on the method of moments.

Table 4. 5: The number of Claims Frequency distribution under Poisson & Geometric distributions of method of moment

No. of Claims	Observed Freq.	Expected Freq. (Poisson)	Expected Freq. (Geometric)
0	643953	642907	643783
1	32178	34176.3	32497.2
2	1784	908.389	1640.41
3	82	16.096	82.8
4	7	0.214	4.18
5	2	0.002	0.211
6	1	2.015	0.011
Total	678007	678008	678008

Table 4. 6: Comparison of absolute deviations of the Poisson & Geometric fit from observed frequencies (Method of Moments)

Number	Observed Claim Freq.	Expected Number of Claims		Absolute Deviations	
		Poisson	Geometric	Poisson	Geometric
0	643953	642907	643783	1046	169.82
1	32178	34176.3	32497.2	1998.3	319.187
2	1784	908.389	1640.41	875.611	143.592
3	82	16.096	82.8	65.904	0.805
4	7	0.214	4.18	65.904	2.82
5	2	0.002	0.211	2.82	1.789
6	1	2.015	0.011	1.789	0.789
Total	**678007**	**678008**	**678008**	**4056.33**	**638.802**
Mean Deviations				**579.475**	**91.257**

It can be inferred from the above results that the two methods of estimating the parameters of the Poisson and Geometric distributions produced approximately the same results. However, the Geometric distribution is found to produce the best fit to the observed number of claim frequencies.

4.4 Analysis on Transition Matrices for NCD Systems

In this section, analysis on the transition matrices and graphs for the hypothetical cases on 3-Level, 4-Level and 5-Level NCD systems are presented. The previous section revealed that Geometric distribution produced the best fit to the claim frequency data. Therefore, the transitional probabilities for the matrices of three hypothetical cases for each 3-Level, 4-Level and 5-Level NCD systems will be computed using the Geometric distribution. The estimated mean parameter for the fitted Geometric distribution was computed as 18.87 under the maximum likelihood estimation method.

The transition matrices for the various hypothetical cases in the 3-Level, 4-Level and 5-Level NCD systems are presented. The transitional probabilities known as the movements of policyholders across the various discount levels are estimated and analysed based on the Geometric distribution.

4.5.1 Results on the 3-Level NCD Systems

The results and discussion on the three cases of the 3-Level discount systems for the movement of policyholders across the discount levels. Figure 4.2 displays the transition matrix and graph for the Case 1, ($\boldsymbol{T_{3,1}}$) of 3-Level discount system. The transition matrix ($\boldsymbol{T_{3,1}}$) shows that there is the likelihood that a policyholder who made claimant at a level would drop to the immediate next level with a probability of 0.048 in the next cycle or year. Alternatively, a policyholder who make claimant at Level 3 or Level 2 will drop to Level 2 or Level 1, respectively with an equal chance of 0.048. However, the chances that a policyholder who do not make any claimant at Level 1 or Level 2 would move to the next level with a probability of 0.952 (i.e. from Level 1 to Level 2 or from Level 2 to Level 3). It is also observed that the likelihood that a policyholder at Level 1, will remain at Level 1 is 0.048 in the next cycle and the likelihood that a policyholder at Level 3 who does not make claim, would remain at the same Level 3 is 0.952. There is no chance that a policyholder not making claim at level 2 will remained at level 2.

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 \\ 0.048 & 0.000 & 0.952 \\ 0.000 & 0.048 & 0.952 \end{bmatrix}$$

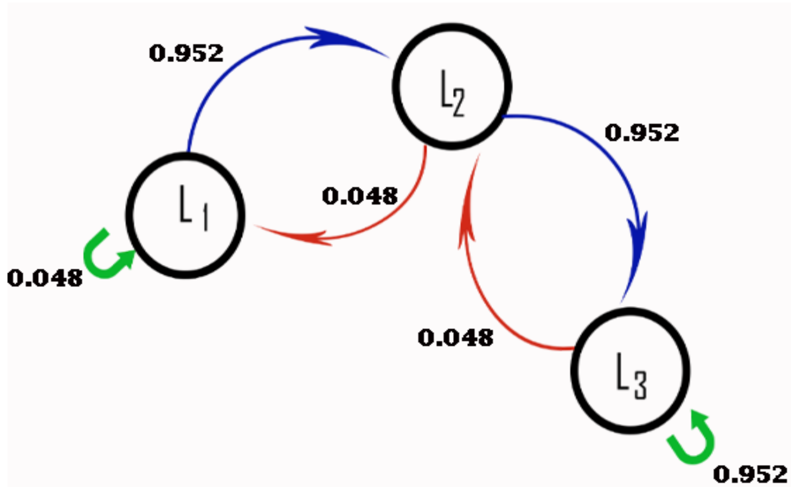


Figure 4. 2: Transition matrix and graph for case 1of 3-Level NCD system ($T_{3,1}$) .

Figure 4.3 presents the transition matrix and graph for the Case 2, ($\boldsymbol{T_{3,2}}$) of the 3-Level discount system. The transition matrix ($\boldsymbol{T_{3,2}}$) and graph show that there is the likelihood that a policyholder who do not make claim would move one level upward with an equal probability of 0.952. However, the chances that a policyholder who make claim at Level 3, may dropped to Level 1 with a probability of 0.05 and the likelihood that a policyholder who make claim at Level 2 will drop to level 1 is 0.048. It is also observed that the likelihood that a policyholder at Level 1 and will remain at the same Level 1 without making claim is 0.048. However, the likelihood that a policyholder at Level 3 who do not make a claim and will remain at the same Level 3 is 0.95. It is observed that there is no chance that a policyholder at Level 2 and will remained at the same Level 2 in the next cycle (Figure 3).

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 \\ 0.048 & 0.000 & 0.952 \\ 0.050 & 0.000 & 0.950 \end{bmatrix}$$

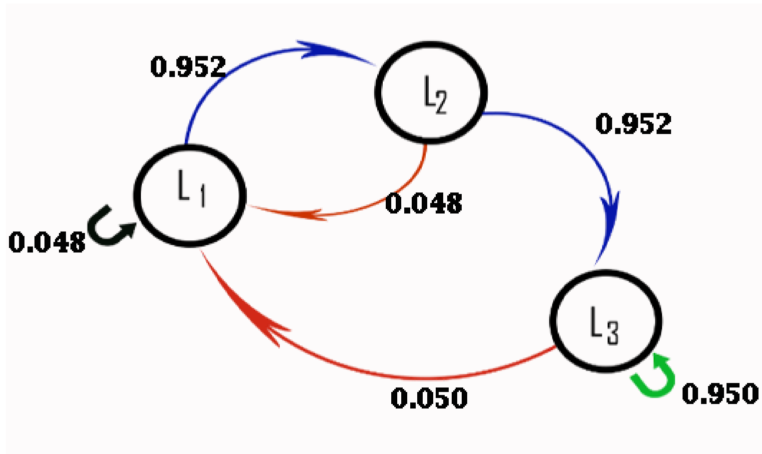


Figure 4. 3: Transition matrix and graph for case 2 of 3-Level NCD system($T_{3,2}$).

The Case 3 of 3-Level NCD system has an additional level known as No Premium level and its transition matrix and graph are displayed in Figure 4.4. It is observed that, Case 3 of 3-Level discount system have a similar movement pattern to that of the Case 1. The only difference is the No Premium level. However, at Level 3, a policyholder who do not make claim would move to the No Premium level for one year before returning to Level 3 again if no claim is made during that one-year period. The probability that a policyholder without making claim at Level 1 or Level 3 will remain in that level in the next cycle are 0.048 and 0.002, respectively.

However, a policyholder who does not made a claim at Level 1 has 0.952 chances of moving up to level 2 and 0.048 likelihood of moving from Level 2 to Level 1 if a claim is made. Again, a policyholder without claim at Level 2 has 0.950 chances of moving up to Level 3 and 0.048 likelihood of dropping from Level 3 to Level 2 if a claim is made.

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 & 0.000 \\ 0.048 & 0.002 & 0.950 & 0.000 \\ 0.000 & 0.048 & 0.002 & 0.950 \\ 0.000 & 0.000 & 1.000 & 0.000 \end{bmatrix}$$

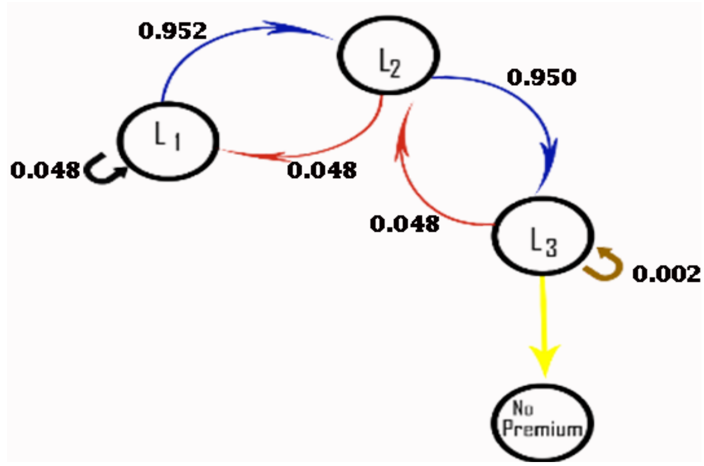


Figure 4. 4: Transition matrix and graph for case 3 of 3-Level NCD system $(T_{3,3})$ **.**

4.5.2 Results on the 4-Level Discount Systems

In this subsection the results, and discussions on the three cases of the 4-Level discount systems for the movement of policyholders across the discount levels are presented. Figure 4.5 displays the transition matrix and graph for the Case 1, $(\boldsymbol{T_{4,1}})$ of 4-Level discount system. The transition matrix $(\boldsymbol{T_{4,1}})$ shows that there is equal probability of a policyholder moving one step upward to the next level if no claim is made at that level and one step downward if a claim is made. It is found that the probability of a policyholder moving one level step upward is 0.952 and the probability of moving one level step downward is 0.048. However, the likelihood that a policyholder without making a claim at Level 1 or Level 4 would remain in that same Level next cycle are 0.048 and 0.952, respectively. There is no likelihood that a policyholder not making claim in Level 2 or Level 3 will remain at that same level the next cycle.

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 & 0.000 \\ 0.048 & 0.000 & 0.952 & 0.000 \\ 0.000 & 0.048 & 0.000 & 0.952 \\ 0.000 & 0.000 & 0.048 & 0.952 \end{bmatrix}$$

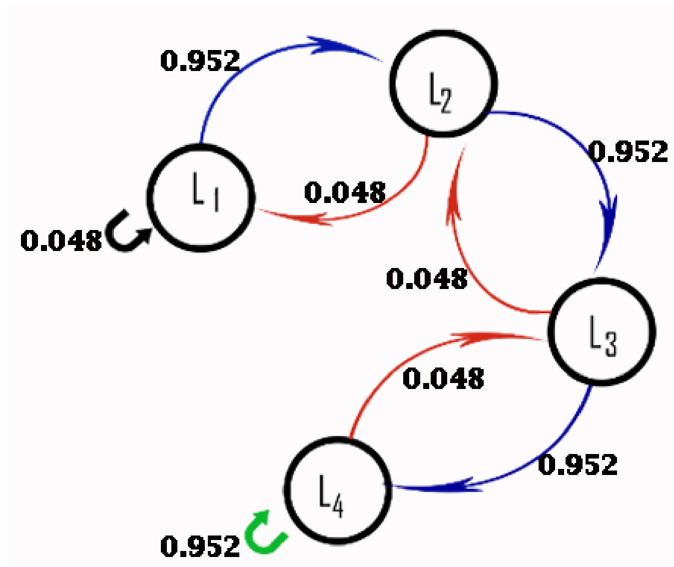


Figure 4. 5: Transition matrix and graph for case 1 of 4-Level NCD System ($T_{4,1}$).

Furthermore, Figure 4.6 presents the transition matrix and graph for the Case 2, ($\boldsymbol{T}_{4,2}$) of 4-Level discount system. The transition matrix ($\boldsymbol{T}_{4,2}$) and graph show that there is the likelihood that a policyholder may move one level step upward with an equal probability of 0.952 in the preceding year if no claim is made and a one level step downward with an equal probability of 0.048 in the preceding year if a claim is made. It is observed that a policyholder making a claim at Level 3 may move to Level 1 with a probability of 0.002 and a policyholder making a claim at Level 4 will move to Level 2 with a probability of 0.002 in the preceding year. However, Case 2 of 4-Level NCD system, has the same chances for movement of policyholders in Case 1 of 4-level NCD system, where a policyholder at Level 1 or Level 4 may remain in that same level in the preceding year with a probability of 0.048 or 0.952, respectively for not making a claim in the current year. It is interesting to note that, there is zero chance that a policyholder not making claim at Level 2 or Level 3 would remain at the same level in the preceding year (Figure 4.6).

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 & 0.000 \\ 0.048 & 0.000 & 0.952 & 0.000 \\ 0.002 & 0.048 & 0.000 & 0.950 \\ 0.000 & 0.002 & 0.048 & 0.950 \end{bmatrix}$$

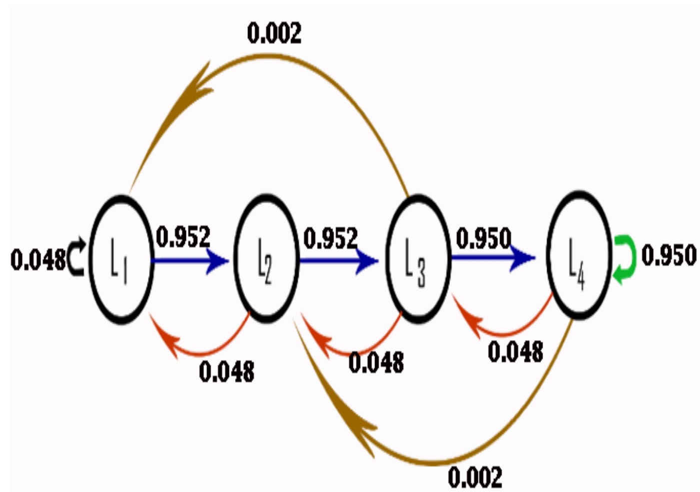


Figure 4. 6: Transition matrix and graph for Case 2 of 4-Level NCD system ($T_{4,2}$).

The Case 3 of 4-Level no claim discount system has an additional level known as No Premium level like the third case of 3-Level NCD system. The transition matrix and graph for this case are displayed in Figure 4.7. It is observed that, Case 3 of 4-Level discount system have a similar movement pattern to that of the Case 2 of 4-Level. The only difference is the No Premium level. However, at Level 4 a policyholder who do not make claim would move to the No Premium level for one year before returning to Level 4 again if no claim is made during that one-year period of no premium payment. The probability that a policyholder at Level 4 without claimant will move to No Premium level is 0.950. Secondly, the probability that a policyholder with no claim at Level 1 will remain in that level in the next cycle is 0.048.

Furthermore, a policyholder who does not make a claim at Level 1, Level 2 or Level 3 has 0.952, 0.952 and 0.950 chances of moving up to Level 2, Level 3 or Level 4, respectively in the next cycle. It is observed that a policyholder making a claim at Level 3 may move to Level 1 with a probability of 0.002 and a policyholder making a claim at Level 4 will move to Level 2 with a probability of 0.002 in the preceding year. Finally, a policyholder with claimant at Level 2, Level 3 or Level 4 will drop to the next level with a probability of 0.048.

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 & 0.000 & 0.000 \\ 0.048 & 0.000 & 0.952 & 0.000 & 0.000 \\ 0.002 & 0.048 & 0.000 & 0.950 & 0.000 \\ 0.000 & 0.002 & 0.048 & 0.000 & 0.950 \\ 0.000 & 0.000 & 0.000 & 1.000 & 0.000 \end{bmatrix}$$

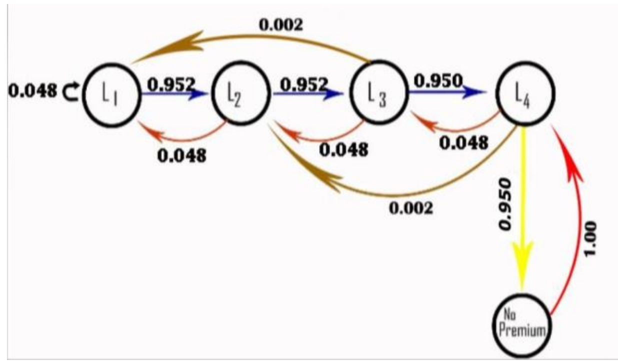


Figure 4. 7: Transition matrix and graph for Case 2 of 4-Level NCD system ($T_{4,3}$).

4.5.3 Results on the 5-Level NCD Systems

This subsection presents the results discussions on the three cases of the 5-Level discount systems for the movement of policyholders across the discount levels. The Markov chain transition matrices and diagrams are presented in Figures 4.8-4.10.

Figure 4.8 displays the transition matrix and graph for the Case 1, ($T_{5,1}$) of 5-Level discount system. The transition matrix ($T_{5,1}$) shows that there are equal probabilities of a policyholder with no claim at a particular moving one step upward to the next level and one step downward to the next level if a claim is made. It is found that the probability of a policyholder moving one level step upward is 0.952 and the probability of moving one level step downward is 0.048. However, the likelihood that a policyholder with no claim at Level 1 or Level 4 would remain in that same Level the next cycle are 0.048 and 0.952, respectively. There is no likelihood that a policyholder not making claim in Level 2 or Level 3 will remain at that same level the next cycle.

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 & 0.000 & 0.000 \\ 0.048 & 0.000 & 0.952 & 0.000 & 0.000 \\ 0.000 & 0.048 & 0.000 & 0.952 & 0.000 \\ 0.000 & 0.000 & 0.048 & 0.000 & 0.952 \\ 0.000 & 0.000 & 0.000 & 0.048 & 0.952 \end{bmatrix}$$

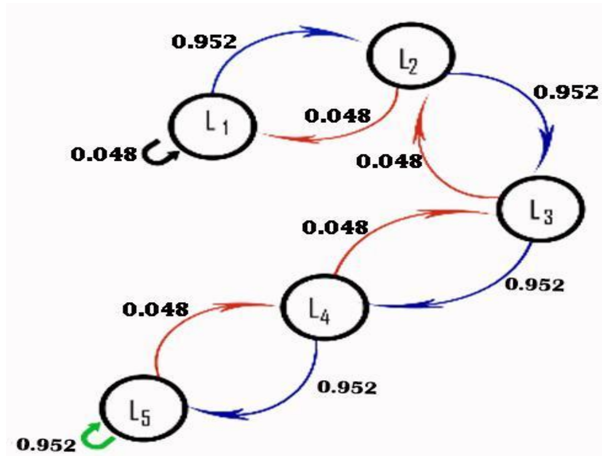


Figure 4. 8: Transition matrix and graph for Case 1 of 5-Level NCD system ($T_{5,1}$).

The results and discussions for the Case 2 of the 5-Level no claim discount system are presented. Figure 4.9 displays the transition matrix and graph for the Case 2 of 5-Level discount system. The transition matrix ($T_{5,2}$) shows that there is an equal probability (0.952) of a policyholder with no claim at Level 1 and Level 2 moving one step upward to the next level (Level 2 and Level 3, respectively) and 0.950 probability of moving from Level 3 to Level 4 and from Level 4 to Level 5 if the policyholder make no claim. However, a policyholder will drop one step to the next level if a claim is made with a probability of 0.048. Furthermore, the likelihood that a policyholder would drop two steps or Levels is 0.002 if a claimant is made (Figure 4.9).

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 & 0.000 & 0.000 \\ 0.048 & 0.000 & 0.952 & 0.000 & 0.000 \\ 0.002 & 0.048 & 0.000 & 0.950 & 0.000 \\ 0.000 & 0.002 & 0.048 & 0.000 & 0.950 \\ 0.000 & 0.000 & 0.002 & 0.048 & 0.950 \end{bmatrix}$$

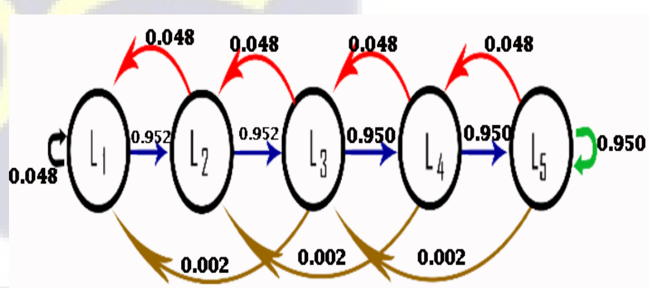


Figure 4. 9: Transition matrix and graph for Case 2 of 5-Level NCD system ($T_{5,2}$)

The final case of 5-Level no claim discount system has an additional level known as No Premium level like the third cases of 3-Level and 4-Level NCD systems. The transition matrix and graph for this system are displayed in Figure 4.10. It is observed that, Case 3 of 5-Level discount system have a similar movement pattern to that of the Case 2 of 5-Level. The only difference is the No Premium level. However, at Level 5 a policyholder who do not make claim would move to the No Premium level the next discount cycle for one year before returning to Level 5 again if no claim is made during that one-year period of no premium payment. The probability that a policyholder at Level 5 without claimant will move to No Premium level is 0.950. The transition matrix ($\boldsymbol{T_{5,3}}$) shows that there is an equal probability (0.952) of a policyholder with no claim at Level 1 and Level 2 moving one step upward to the next level (Level 2 and Level 3, respectively) and 0.950 probability of moving from Level 3 to Level 4 and from Level 4 to Level 5 if the policyholder makes no claim. However, a policyholder will drop one step to the next level if a claim is made with a probability of 0.048. Furthermore, the likelihood that a policyholder would drop two steps or Levels is 0.002 if a claimant is made (Figure 4.10).

$$\begin{bmatrix} 0.048 & 0.952 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.048 & 0.000 & 0.952 & 0.000 & 0.000 & 0.000 \\ 0.002 & 0.048 & 0.000 & 0.950 & 0.000 & 0.000 \\ 0.000 & 0.002 & 0.048 & 0.000 & 0.950 & 0.000 \\ 0.000 & 0.000 & 0.002 & 0.048 & 0.000 & 0.950 \\ 0.000 & 0.000 & 0.000 & 0.000 & 1.000 & 0.000 \end{bmatrix}$$

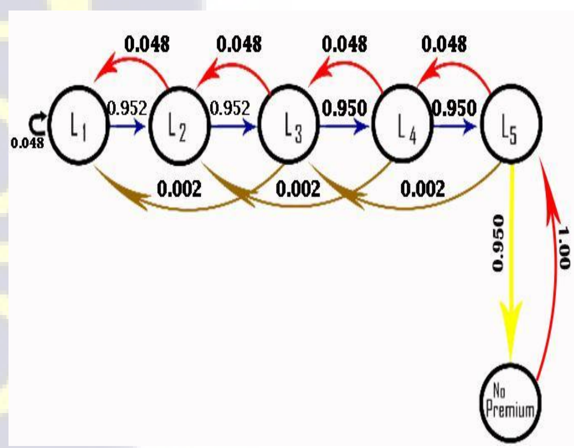


Figure 4. 10: Transition matrix and graph for Case 1 of 5-Level NCD system ($\boldsymbol{T_{5,3}}$).

The results observed from Figures 4.2, 4.5 and 4.8 show that the system treats policyholders fairly. It is generally observed that, policyholders have a higher chance of moving up to the next level if claims are not made and the probability of dropping one level if claims are made is very small. The system penalizes claimants by reducing their discount one level at a time or maintaining their discount level if they are already at the least level of discount. Non-claimants are equally rewarded one level at a time.

The cases 2 of the three NCD systems showed a more penalized and strict system. For instant, Case 2 of 3-Level discount system penalized policyholders with a probability of 0.05 of losing the highest discount-level and back to Level 1 to pay full premiums and that can also be seen 4-Level and 5-Level at a probability of 0.002. On the contrary, the event of losing more than one discount-level at a time, was impossible in the other NCD systems.

Finally, the third case of the three NCD systems are however ideal systems. There are high chances of 0.950 moving to the no premium level if claims are not made. This encourages policyholders not to make claims when minor accidents occur. Generally, the no claim systems are designed to reward policyholders who are extra vigilant on the road in terms of avoiding road accidents and penalized bad road users. Therefore, claimants are grouped into 2: those at-fault and those not-at-fault. Those at-fault depending upon the magnitude of the accident and claim are moved one or two discount-level down and those who are not-at-fault are saved or spared any penalties; they are retained at their discount-level just before the claim.

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter presents summary, conclusion and recommendations of the study based on the findings of the study. The study employed maximum likelihood and moment methods of estimation to fit the mixed Poisson-exponential model and the Markov Chains technique was applied to calculate the probabilities for the three hypothetical No-Claims Discount model systems.

5.2 Summary

The study seeks to model a multi-level discount system that analyses the problem of homogeneity in the treatment of motor insurance policyholders in a typical No-Claims Discount system. The summary of the findings based on the set objectives are as follows.

The descriptive statistics of the claim frequency revealed that the number of claims by policyholders was right or positively skewed with a mean number of claims approximately zero (0.053). The results revealed that the Geometric distribution was better fitted to the observed claim frequencies based on the maximum likelihood estimation method than the Poisson distribution. Also, under the method of moment estimation, the Geometric distribution provides a much better fit to the observed claim frequencies than the Poisson distribution. It was observed that the geometric model performed approximately the same under both methods of estimating the parameters.

The transition matrices for the various hypothetical cases in the 3-Level, 4-Level and 5-Level NCD systems were estimated using the geometric probability model based on the maximum likelihood estimation method. the results for the 3-Level NCD systems showed that the policyholders were rewarded with approximately 95% chances of moving to the next higher level towards attaining a no premium zone if a claim is not made in the cycle and were punished if a claim is made by dropping to the lower level with approximate probability of 0.048.

For the 4-Level and 5-Level NCD systems, there was not much difference with the 3-Level NCD systems. With the 4-Level and 5-Level NCD systems, a policyholder maybe punished with one-two levels drop from the current level if a claim is made and one step upward if no claim is made as a reward. Generally, there was approximately 0.950 probability of a policyholder moving to a higher level if a no claim is made during that period and 0.048 probability of dropping to a lower level if a claim is made.

5.3 Conclusion

The study's aim was to model motor insurance claims of policyholders using Markov chains and mixed Poisson-exponential model. The claim count distribution was found to follow a Geometric distribution and the maximum likelihood estimation and method of moments produced approximation same results for estimating the parameters of the Geometric model (distribution).

The results revealed that the system treats the policyholders fairly with higher chances rewards for not making claims in any given level. However, due to the fairness of the system, policyholders who make claims are equally punished by dropping from their current level to the lower one or are made to stay in their current level in the next cycle. The second cases of the three systems were considered strict and the third cases were considered ideal systems. This because policyholders were given high chances of moving to the no premium level if claims are not made. Therefore, policyholders are encouraged not to make claims when minor accidents occur when they are at the third case system. In conclusion, the multi-level NCD system was designed to reward policyholders who are extra vigilant on the road in terms of avoiding road accidents and penalized bad road users.

5.4 Recommendations

The study presents the following recommendations to stakeholders in the insurance industry and researchers based on the findings of the study.

1. The motor insurance institutions in Ghana should incorporate No-Claims Discount systems in pricing insurance premiums in Ghana, especially comprehensive insurance.
2. Further work should be done on multi-level systems: 6-Level, 7-Level, 8-Level and 9-Level NCD systems to unravel the policyholders' behaviours on road usage.
3. This study considered only the Poisson and exponential mixed distribution as an alternative to the Poisson distribution. Future studies could consider the Poisson-Lindley, negative Binomial and the one-parameter or two-parameter Gamma distributions and Geometric and two parameter gamma distributions.

REFERENCES

Anyango, O. B. (2015). *Markov chains and mixed Poisson distributions in the No Claims Discount Systems*.

Baione, F., Levantesi, S., & Menzietti, M. (2002). The Development of an Optimal Bonus-Malus System in a Competitive Market. *ASTIN Bulletin*, *32*(1), 159–170. https://doi.org/10.2143/ast.32.1.1021

Burney, S. M. A., Khan, L. M., Burney, S., & Humayoun, M. (2022). Data Analysis and Modelling of Claim Amounts of Car Insurance using Big Data: A Study for Pakistan. *Asian Journal of Probability and Statistics, Vol. 19(4):* pp:46-53. DOI: 10.9734/AJPAS/2022/v19i430476

Chehui, Zhangjiwu, & Zhangxingyang. (2011). Research on motor vehicle insurance underwriting risk management model. *Procedia Engineering*, *15*, 4973–4977. https://doi.org/10.1016/j.proeng.2011.08.924

Frangos, N. E., & Vrontos, S. D. (2001). Design of Optimal Bonus-Malus Systems With a Frequency and a Severity Component On an Individual Basis in Automobile Insurance. *ASTIN Bulletin*, *31*(1), 1–22. https://doi.org/10.2143/ast.31.1.991

Frees, E. W., & Valdez, E. A. (2008). Hierarchical insurance claims modeling. *Journal of the American Statistical Association*, *103*(484), 1457–1469. https://doi.org/10.1198/016214508000000823

Heras, A., Gil, J. A., García-Pineda, P., & Vilar, J. L. (2004). An Application of Linear Programming to Bonus Malus System Design. *ASTIN Bulletin*, *34*(2), 435–456. https://doi.org/10.1017/S0515036100013775

Hosgood, G. L. (n.d.). *LSU Digital Commons Markov models to estimate and describe survival* *time and experience in cohorts with high euthanasia frequency Recommended Citation*. https://digitalcommons.lsu.edu/gradschool_dissertations

Jacob, A., & Wu, Z. (2020). An Alternative Pricing System through Bayesian Estimates and Method of Moments in a Bonus-Malus Framework for the Ghanaian Auto Insurance Market. *Journal of Risk and Financial Management*, *13*(7), 143. https://doi.org/10.3390/jrfm13070143

Janczura, J. & Weron, R. (2010). An empirical comparison of alternate regime-switching models or electricity spot prices. *Munich Personal RePEc Archive (MPRA)*,pp: 1-21. Available: https://mpra.ub.uni-muenchen.de/id/eprint/20546

Karlis, D., & Xekalaki, E. (2005). Mixed Poisson Distributions. *International Statistical Review / Revue Internationale De Statistique*, *73*(1), 35-58. Retrieved June 25, 2021, from http://www.jstor.org/stable/25472639

Kliger, D., & Levikson, B. (2002). Pricing no claims discount systems. In *Insurance: Mathematics and Economics* (Vol. 31).

Kumar, R., Singh, A. K., Goh, K. L., Goh, A., & Leng, K. (2011). *Application of Markov Chain in the PageRank Algorithm Technical reports View project 7th International Conference on Smart Computing & Communications (ICSCC) @ Sarawak Malaysia View project Application of Markov Chain in the PageRank Algorithm*. https://www.researchgate.net/publication/256089617

Levin, D. A., Peres, Y., & Wilmer, E. L. (n.d.). *Markov Chains and Mixing Times*.

Manan, N. B., Hashim, H., & Mohd, M. A. (2013). Modification of the current Malaysian No-Claim Discount system using Markov Chains. *BEIAC 2013 - 2013 IEEE Business Engineering and Industrial Applications Colloquium*, 926–931.

https://doi.org/10.1109/BEIAC.2013.6560273

Markov Chains. ***Brilliant.org***. Retrieved 09:42, June 24, 2021, from https://brilliant.org/wiki/markov-chains/

Mutiu, S., & Dotun, O. (n.d.). *Application of Markov Chain in Forecasting: A Study of Customers' Brand Loyalty for Mobile Phones*. www.euacademic.org

Myers, D. S., Wallin, L., & Wikström, P. (n.d.). *MVE220 Financial Risk: Reading Project An introduction to Markov chains and their applications within finance Group Members*.

Nath, D. C. & Sinha, P. (2014). A Markovian study of no claim discount system of Insurance Regulatory and Development Authority and its application. *Thailand Statistician, Vol. 12(2), pp: 223-236* http://statassoc.or.th *Operations Research (Hillier & Lieberman)*.

Petrović, G., Petrović, N., & Marinković, Z. (2008). APPLICATION OF THE MARKOV THEORY TO QUEUING NETWORKS *. In *FACTA UNIVERSITATIS Series: Mechanical Engineering* (Vol. 6, Issue 1).

Pfeifer, P. E., & Carraway, R. L. (2000). MODELING CUSTOMER RELATIONSHIPS AS MARKOV CHAINS. In *Journal of Interactive Marketing* (Vol. 14, Issue 2).

Pitrebois, S., Walhin, J. F., & Denuit, M. (2005). Bonus-malus Systems with Varying Deductibles. *ASTIN Bulletin*, *35*(1), 261–274. https://doi.org/10.1017/S051503610001415X

Samaniego, J. M. (2003). *M A R K O V C H A I N S*.

Sonnenberg, F. A., & Beck, J. R. (2009). *at National Institutes of Health Library on*. http://mdm.sagepub.com

Walhin, J. K. & Paris, J. (1999). Using Mixed Poisson Processes in Connection with Bonus-Malus Systems. *ASTIN Bulletin*, *29*(1), 81–99. https://doi.org/10.2143/ast.29.1.504607

Willmot G.E., Lin X.S. (2001) Mixed Poisson distributions. In: Lundberg Approximations for Compound Distributions with Insurance Applications. Lecture Notes in Statistics, vol 156. Springer, New York, NY. https://doi.org/10.1007/978-1-4613-0111-0_3

Wu, X., Chen, M., Guo, J., & Jin, C. (2015). On a discrete-time risk model with claim correlated premiums. *Annals of Actuarial Science*, *9*(2), 322–342. https://doi.org/10.1017/s1748499515000032