MARKOV CHAINS AND MIXED POISSON DISTRIBUTIONS IN THE NO CLAIMS DISCOUNT SYSTEMS

BY

OGENO BEATRICE ANYANGO

A RESEARCH REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE AWARD OF A MASTER OF SCIENCE DEGREE IN ACTUARIAL SCIENCE OF THE UNIVERSITY OF NAIROBI.

JUNE 2015

DECLARATION

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

Ogeno Beatrice Anyango **Date**

This research project has been submitted for presentation with my approval as the university supervisor:

Prof. J.A.M. Ottieno **Date**

ACKNOWLEGMENT

To the Almighty Lord, who has guided me all through, I thank you Lord for bringing me this far and guiding me through my studies.

My gratitude also goes to my supervisor, Prof. J.A.M. Ottieno, for his support, experience and encouragement that led to the successful completion of this project.

I want to appreciate my family and friends for the support and encouragement that they gave me throughout this period.

I would also want to show my appreciation to my colleagues as well, to whom Iôm grateful for their assistance throughout my studies.

DEDICATION

To my parents, Jackton O. Ogeno and Salome A. Ogeno, and the entire family for their support they gave me. God bless you all.

ABSTRACT

Most NCD systems are unfair to either or both parties. Most 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 the extreme cases, whereby if a driver makes claim(s), he loses all the discounts accumulated over the years and goes back to the level of full premium payment. The other movements within the NCD systems are those of the in-between cases, these consider the frequency of claims, and can involve moving few steps back the discount level in case of claim(s). A fair NCD system, should take into consideration the frequency of claims and the non-homogeneity factor.

In this paper, we have used the Markov chains to explain the movement between levels and used the Mixed Poisson distributions to calculate probabilities, with the mixing distributions being the exponential distribution, the one parameter gamma distribution, the two parameter gamma distribution and the Lindley distribution. The rules applied in different systems have been analysed and combined to take into consideration the claims frequency.

The following distributions have been constructed, the Geometric, the Negative Binomial distribution with one and two parameters, and the Poisson Lindley, and their parameters estimated using the method of moments and the maximum likelihood method. The distributions have then been used to fit the claim frequency data and a comparison made with data from a Poisson ï Inverse Gaussian distribution.

TABLE OF CONTENTS

DECLARATION ... **ii**

ACKNOWLEGMENT ... **iii**

DEDICATION ... **iv**

TABLE OF CONTENTS ... **vi**

CHAPTER I ... **1**

GENERAL INTRODUCTION ... **1**

1.1 BACKGROUND INFORMATION ... 1

THE EFFECTS OF THE NO CLAIMS DISCOUNT SYSTEMS ... 2

1.2 STATEMENT OF THE PROBLEM ... 2

1.3 OBJECTIVES OF THE STUDY ... 3

1.4 LITERATURE REVIEW ... 4

1.5 SIGNIFICANCE OF THE STUDY ... 5

CHAPTER II ... **6**

2.1 INTRODUCTION ... 6

2.2 MARKOV CHAINS ... 6

2.2.1 Definition ... 6

2.2.2 Conditional probability ... 6

2.2.3 Notations and terminologies ... 7

2.2.4 Higher orders of transition probabilities ... 8

2.2.5 Classification of states ... 8

2.2.6 Classification of the Markov chains ... 13

2.2.7 Invariant (stationary) distribution ... 14

CHAPTER III ... **15**

NO CLAIMS DISCOUNT SYSTEMS ... **15**

3.1 INTRODUCTION ... 15

3.2 THE THREE STAGE BMS ... 15

3.3. THE FOUR STAGE NCD ... 22

3.4 THE FIVE STAGE BMS ... 26

3.5 THE SIX STAGE BMS ... 30

3.6 THE SEVEN STAGE BMS ... 42

3.4. THE EIGHT STAGE NCD SYSTEM ... 52

3.7 THE NINE STAGE BMS ... 53

3.9. THE TEN STAGE BMS 57
CHAPTER IV 62
CLAIM FREQUENCY DISRIBUTION........ 62
4.1 INTRODUCTION 62
4.2 PARAMETER ESTIMATION 62
4.2.1 THE METHOD OF MOMENTS........ 62
4.2.2 MAXIMUM LIKELIHOOD METHOD 63
4.3 GAMMA FUNCTIONS 64
4.4 MIXED POISSON DISTRIBUTIONS........ 66
4.4.1 INTRODUCTION........ 66
4.4.2 POISSON ï EXPONENTIAL MIXTURE (GEOMETRIC DISTRIBUTION)..... 68
4.4.3 POISSON-GAMMA MIXTURE (NEGATIVE BINOMIAL DISTRIBUTION).. 71
4.4.3.1 The one parameter gamma mixing distribution 71
4.4.3.2 The two parameter gamma mixing distribution 75
4.5 THE POISSON LINDLEY DISTRIBUTION 79
CHAPTER V 84
DATA ANALYSIS........ 84
5.1 INTRODUCTION 84
5.1.1 Method of moments........ 84
5.1.2 Maximum likelihood method........ 86
5.2 REMARKS AND RECOMMENDATIONS 89
REFERENCES........ 90

CHAPTER I

GENERAL INTRODUCTION

1.1 BACKGROUND INFORMATION

It is quite common in general insurance, for instance in the automobile insurance, for premium to be reduced by a certain factor, by the insurer, for a policyholder who has made no claim in a given period. In case of a claim, the premium is increased, also by a certain factor (a process known as loading).

Either system amount to what is known as a multi-layer premium system, also known as the no-claims-discount system (NCD system) or the Bonus Malus System (BMS).

The BMS are used by insurers to categorize policyholders into homogeneous risk groups that pay premiums relative to their claims experience. Once categorized, the risks can be rated using generalized linear models. Methods used in determination of fair premium that reflects the individual risk of a driver include:

1. The use of variables such as policyholders age, occupation, gender, degree of disability, the type and use of car, and the place of garage. These variables are used in dividing automobile risks into different homogeneous classes, a method known as the priori rating.
2. Policyholders from a given risk cell are subdivided into bonus-malus classes and their claims histories. This use of individual past claim history is known as the posterior rating, for example the BMS.

NCD is determined by three elements: the premium scale, initial class and the transition rules. These determine the transfer from one class to another when the number of claims is known. An insured enters the system in the initial class when he applies for insurance and throughout the entire driving lifetime, the transition rules are applied upon each renewal to determine the new class.

Depending on rules, new policyholders may be required to pay full premium initially and obtain discounts in future as a result of the claim free years.

THE EFFECTS OF THE NO CLAIMS DISCOUNT SYSTEMS

This scheme entices the insured who makes few claims in the recent years to stay with the company by rewarding him with discounts on initial premium.

In general, it rewards policyholders for not making claims during a year, i.e. granting bonuses to careful drivers. This leads to the existence of the phenomenon, hunger for bonus (Philipson C., 1960). It also reduces the cost of insurance cover and discourages small claims.

Through NCD, drivers become more responsible about their vehicles and when driving. Drivers can also insure their NCD to protect it, once maximum discount is reached. The charge is added to the insurance policy.

One of the features of a NCD system is its ability to deal with the problem of adverse selection as it is designed to evaluate true distribution of reported accidents relating to unchanging characteristics.

Another role is linked to moral hazard. Distribution of reported accidents over time must be taken into account to maintain the incentives for cautious behaviour at an optimal level (Dionne, 2005).

Dionne and Dostie (2007), show that such a system has two effects when the insurance industry is committed to its application. First, since past claims are associated with increased insurance premiums in the future (moral hazard), motivation of drivers will be more prudent. Secondly, there will be an improvement in risk classification by allowing insurance companies to make a bad risk pay more and a good one to pay less.

1.2 STATEMENT OF THE PROBLEM

An optimal NCD system should be efficient and competitive, aiming to relate as best as possible the premium paid by the insured to his driving experience (Guerreiro and Mexia, 2002). This is not the case in most countries.

These systems are established through the free market and others are government imposed. The free market involves the insurance companies coming up with their own systems while another way is where the government introduces the rules to be used by the insurance companies. The effectiveness of these NCD systems has been doubted for a long time.

Most systems treat policyholders as homogeneous, which is not usually the case.

The Kenyan system has seven classes, with levels, 40, 50, 60, 70, 80, 90 and 100. The merit factors ranges from 1 (coefficient for the entry class) to 0.4. The problem with this system is no potential policyholder may expect it to be transparent as no provision seems to have been made in the premium construction to accommodate the bonus granted to the good drivers (Verico 2002).

In UK, the highest level (50 ï 60) is reached after 4 to 6 years making it a marketing scheme rather than a way of distinguishing risk.

In Germany, it takes up to 25 years to earn the comparative level of bonus (Schmitt, 2000).

In Brazil, motor insurance policyholders are subdivided into 7 classes with premium levels, 100, 90, 85, 80, 75, 70 and 65, where newcomers join in class 7, at level 100. Each claim-free year results in a one-class discount and each at-fault claim is penalized by one class. Simple system, but it is doubtful if it will effectively motivate safe driving.

Most drivers in Portugal escape the malus due to lack of efficient transfer of information between insurers. After making a claim the policyholder would leave his insurer and buy another policy, from a competitor (Guerreiro and Mexia, 2002).

The motive for this study is the lack of efficiency in the BMSs in use. They encourage non-reporting of minor claims rather than safe driving, they do not differentiate between small and big size claims, which is unfair and it is more of a marketing scheme and they do not take into consideration the distinction of the risks.

In Kenya, most policyholders take the automobile policies that only cover the third party, and rarely take into consideration the comprehensive cover as they consider this too expensive. Even with the NCD in place they still do not take cover for their own vehicles.

1.3 OBJECTIVES OF THE STUDY

The main objective

- To evaluate the different multi-layer systems and to consider risk distinction.

The specific objective

- To consider the different stages of the NCD systems and the rules applied.
- To combine different transition rules for each stage so as to come up with a system that is fair to both parties.
- To calculate the claim frequency distribution using the mixed Poisson distribution.

1.4 LITERATURE REVIEW

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 is the government imposed systems, where the rules that are applied in the NCD systems are governed by the government and every insurer has to apply these rules.

What follows is the lit review on different rules applied in both types of systems, used by different countries.

Das and Basu (2003) use two systems in a 3 by 3 stage NCD system that show two possible scenarios in terms of movement by drivers between the various levels of discount, with the assumption that the claim amount has a log-normal distribution. The systems do not fairly share the cost of premium between good and bad drivers.

Institute and Faculty of Actuaries Examination, CT 4 (2009), over the years has several questions with the NCD systems with different rules applied within these systems. From the 3 by 3 stage NCD systems to the 4 by 4 NCD systems, most of which are assumed to follow a Poisson distribution.

Lemaire and Zi (1994), 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 policyholdersô 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.

Emamverdi 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.

Nath and Sinha (2014), inspect the desirability of the multi-layer premium system of the Insurance 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. The levels of the NCD should be adjusted, so that bad drivers pay double premium that of good drivers.

Ibiwoye et al (2011), evaluate the 6 by 6 stage bonus-malus system in practice in Nigeria, and they observe that the system is far from optimal due to a number of weaknesses, for instance it does not consider frequency of claims, for a claim reported the policyholderøs bonuses are cancelled, among others. They then construct a scale that has reasonable penalties and is commercially feasible.

Soren Asmussen (2014), inspects how reasonable it is to view BMS via the stationary distribution. Among the systems he looks at is the one in use in Ireland, the bonus rules applied in the Irish system with six levels. His conclusion is that, transient distributions are far from the stationary, and this has considerable consequences on computation of average premiums.

Jean Lemaire (1998) looks at the merit-rating technique used in most of Europe and Asia. One of the countries he considers is Brazil, with seven levels. The distribution of the number of claims is assumed to conform to a Poisson distribution. He discovers that rating freedom encourages insures to adapt tougher systems, and that most companies are now not using the same NCD systems with deregulation ideas gaining ground.

Walhin and Paris (1999), clarify on how to construct a bonus-malus table using the principle of zero utility when working with mixed Poisson distributions. They use a nine by nine stage system as an illustration, and they discover that the parametric mixed Poisson distribution they use is a slightly better fit than the negative binomial or the Poisson Inverse Gaussian.

Most 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 the extreme cases, where if a driver makes claim(s), he loses all the discounts accumulated over the years and goes back to the level of full premium payment. The other movements within the NCD systems are those of the in-between cases, these consider the frequency of claims, and can involve moving few steps back the discount level in case of claim(s).

1.5 SIGNIFICANCE OF THE STUDY

This study is significant in the automobile insurance as it combines different transition rules for each stage so as to come up with systems that are fair to both parties, by taking into consideration the claims frequency and risk distinction.

Due to hunger for bonus, this will lead to drivers being more careful so as to get more discounts. It will also discourage the reporting of small claims, as most policyholders would not want to lose the discount they have accumulated over time.

CHAPTER II

2.1 INTRODUCTION

The NCD systems in this project are based on the Markov chains. In this chapter, the theories of Markov chains have been discussed.

2.2 MARKOV CHAINS

Modern probability theory studies chance processes for which the knowledge of previous outcomes influences predictions for future experiments. In principle, when we observe a sequence of chance experiments, all of the past outcomes could influence our predictions for the next experiment.

In 1907, A. A. Markov began the study of an important new type of chance process. In this process, the outcome of a given experiment can affect the outcome of the next experiment. This type of process is called a Markov chain.

2.2.1 Definition

Markov chain- is a sequence of random variables X_n, with the Markov property.

A discrete-time stochastic process X_n is said to be a Markov Chain if it has the Markov property.

Markov property – the future event depends on the immediate past and not the remote past. The future evolution of a system depends only on the current state of the system and not the past history.

Random variable ï The mathematical rule (or function) that assigns a given numerical value to each possible outcome of an experiment in the sample space of interest the future outcome is conditionally independent of the past given the present (Memoryless)

2.2.2 Conditional probability

Let us consider two events, E_1 and E_2, by Bayes theorem:

$$\Pr ob(E_2 | E_1) = \frac{\Pr ob(E_2 | E_1)}{\Pr ob(E_1)}$$

$$\therefore \Pr ob(E_1, E_2) = \Pr ob(E_2 | E_1) \Pr ob(E_1)$$

For three events, E_1, E_2 and E_3, we have:

$$\Pr ob(E_3 | E_2, E_1) = \frac{\Pr ob(E_1, E_2, E_3)}{\Pr ob(E_1, E_2)}$$
$$\Pr ob(E_1, E_2, E_3) = \Pr ob(E_3 | E_2, E_1) \Pr ob(E_2 | E_1) \Pr ob(E_1)$$
$$\Pr ob(E_1, E_2, E_3) = \Pr ob(E_3 | E_2) \Pr ob(E_2 | E_1) \Pr ob(E_1)$$

Markov property means that:

$$\Pr ob(E_1, E_2, E_3) = \Pr ob(E_3 | E_2)$$

2.2.3 Notations and terminologies

Events are called states, and a set of events $\{E_1, E_2, E_3,...\}$ is called state space.

$\Pr ob\{E_j | E_i\}$ is called transitional probability P_{ij}

$$P_{ij} = \Pr ob\{E_j | E_i\}$$

=Probability of moving from state E_i to state E_k

=Prob$\{E_i \longrightarrow E_j\}$

Transition probability matrix is given by;

$$P = P_{ij} = \begin{array}{c} \\ E_0 \\ E_1 \\ E_2 \\ E_3 \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccc} E_0 & E_1 & E_2 & E_3 & \cdots \end{array} \\ \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} & \cdots \\ P_{10} & P_{11} & P_{12} & P_{13} & \cdots \\ P_{20} & P_{21} & P_{22} & P_{23} & \cdots \\ P_{30} & P_{31} & P_{32} & P_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \end{array}$$

Where;

$0 \le P_{ij} \le 1$ and $\sum_j P_{ij} = 1$

With the notations above;

$$\begin{aligned} \Pr ob(E_1, E_2, E_3, E_4) &= \Pr ob(E_4 | E_3) \Pr ob(E_3 | E_2) \Pr ob(E_2 | E_1) \Pr ob(E_1) \\ &= \Pr ob(E_1) \Pr ob(E_2 | E_1) \Pr ob(E_3 | E_2) \Pr ob(E_4 | E_3) \\ &= \Pr ob(E_1) P_{12} P_{23} P_{34} \\ &= a_1 P_{12} P_{23} P_{34} \end{aligned}$$

Where $a_1 = \Pr ob(E_1) =$ initial probability

Generally, a Markov chain has two sets of probabilities:

- Initial / absolute probability
- Conditional / transition probability

2.2.4 Higher orders of transition probabilities

Second order of transition matrix is denoted by $P_{ij}^{(2)}$

$P_{ij}^{(2)}$ is the probability of moving from state E_i to state E_j in two steps.

$$P_{ij}^{(2)} = \sum_v P_{iv} P_{vj}$$

$$P^2 = PP$$

Extending to P^3

$$P^3 = PP^2$$

This implies that, $P_{ij}^{(3)} = \sum_v P_{iv} P_{vj}^{(2)}$

Or $P^3 = P^2 P$

$$P_{ij}^{(3)} = \sum_v P_{vj}^{(2)} P_{iv}$$

In general,

$$P^n = P^{n-1} P \Rightarrow P_{ij}^{(n)} = \sum_v P_{vj}^{(n-1)} P_{iv}$$

$$P^n = PP^{n-1} \Rightarrow P_{ij}^{(n)} = \sum_v P_{iv} P_{vj}^{(n-1)}$$

More generally,

$$P^{m+n} = P^m P^n \Rightarrow P_{ij}^{(m+n)} = \sum_v P_{iv}^{(m)} P_{vj}^{(n)}$$

This is called the Chapman Kolmogorov equations or formula.

2.2.5 Classification of states

- **<u>Return probabilities</u>**

Let $f_{jj}^{(n)}$ = the probability of returning to state E_j in n steps for the first time.

And $p_{jj}^{(n)}$ = the probability of returning to state E_j in n steps but not necessarily for the first time.

$f_{jj}^{(n)} \geq p_{jj}^{(n)}$

Assumptions: $f_{jj}^{(0)} = 0$ and $p_{jj}^{(0)} = 1$

Then the relationship between the two will be;

$$p_{jj}^{(n)} = \sum_{v=1}^{n} f_{jj}^{(v)} p_{jj}^{(n-v)} \text{ ; } n \geq 1 \cdots (i)$$

Let $F(S) = \sum_{n=0}^{\infty} f_{jj}^{(n)} S^n$

and $P(S) = \sum_{n=0}^{\infty} p_{jj}^{(n)} S^n$

$$F(S) = f_{jj}^{(0)} + \sum_{n=1}^{\infty} f_{jj}^{(n)} S^n$$

$$F(S) = \sum_{n=1}^{\infty} f_{jj}^{(n)} S^n \cdots (ii)$$

$$P(S) = p_{jj}^{(0)} + \sum_{n=1}^{\infty} p_{jj}^{(n)} S^n$$

$$P(S) = 1 + \sum_{n=1}^{\infty} p_{jj}^{(n)} S^n \cdots (iii)$$

Multiplying equation (i) by S^n and then summing the results over n :

$$\sum_{n=1}^{\infty} P_{jj}^{(n)} S^n = \sum_{n=1}^{\infty} \sum_{v=1}^{n} f_{jj}^{(v)} p_{jj}^{(n-v)} S^n$$

$$= \sum_{n=1}^{\infty} \sum_{v=1}^{n} \left(f_{jj}^{(v)} S^v \right) \left(P_{jj}^{(n-v)} S^{n-v} \right)$$

From equation (iii) we have;

$$P(S) - 1 = \sum_{n=1}^{\infty} p_{jj}^{(n)} S^n$$

$$= \sum_{n=1}^{\infty} \sum_{v=1}^{n} \left(f_{jj}^{(v)} S^v \right) \left(P_{jj}^{(n-v)} S^{n-v} \right)$$

$$= \sum_{n=1}^{\infty} \sum_{v=1}^{\infty} \left(f_{jj}^{(v)} S^v \right) \left(P_{jj}^{(n-v)} S^{n-v} \right)$$

$$= \sum_{v=1}^{\infty} \left\{ f_{jj}^{(v)} S^v \sum_{n=1}^{\infty} P_{jj}^{(n-v)} S^{n-v} \right\}$$

$$= \sum_{v=1}^{\infty} \left\{ f_{jj}^{(v)} S^v \sum_{n=v}^{\infty} P_{jj}^{(n-v)} S^{n-v} \right\}$$

$$= \sum_{v=1}^{\infty} \left\{ f_{jj}^{(v)} S^v \left[p_{jj}^{(0)} + p_{jj}^{(1)} S + p_{jj}^{(2)} S^2 + p_{jj}^{(3)} S^3 + \cdots \right] \right\}$$

$$= \sum_{v=1}^{\infty} \left(f_{jj}^{(v)} S^v \sum_{n=0}^{\infty} p_{jj}^{(n)} S^n \right)$$

$$\therefore P(S) - 1 = \sum_{v=1}^{\infty} f_{jj}^{(v)} S^v P(S)$$

$$= P(S) \sum_{v=1}^{\infty} f_{jj}^{(v)} S^v$$

$$= P(S) F(S)$$

$$P(S) - P(S) \bullet F(S) = 1$$

$$P(S) = \frac{1}{1 - F(S)}$$

- **Persistency, Transiency and Periodicity**

Let $f_j = \sum_{n=1}^{\infty} f_{jj}^{(n)}$

= the probability of eventually returning to state E_j

Persistency – a state E_j is persistent (recurrent) if:

$$f_j = \sum_{n=1}^{\infty} f_{jj}^{(n)} = 1$$

But $0 \le f_{jj}^{(n)} \le 1$ and $\sum_{n=1}^{\infty} f_{jj}^{(n)} = 1$

Thus $\left\{ f_{jj}^{(n)} : n = 1, 2, 3, \cdots \right\}$ is a probability mass function (pmf)

Furthermore, $f_{jj}^{(n)}$ is the probability of returning to E_j in n for the first time. Such a pmf is called first passage distribution, i.e. $\left\{ f_{jj}^{(n)} : n = 1, 2, 3, \cdots \right\}$ is a first passage distribution pmf, with mean usually known as mean recurrence time and denoted by μ_j which is given as; $\mu_j = \sum_{n=1}^{\infty} n f_{jj}^{(n)}$

If $\mu_j = \infty$ then E_j is called null persistent state.

If $\mu_j \lhd \infty$ then E_j is called non-null persistent state.

Transiency – a state E_j is transient if:

$$f_j = \sum_{n=1}^{\infty} f_{jj}^{(n)} \lhd 1$$

Alternatively,

$$\sum_{n=0}^{\infty} p_{jj}^{(n)} S^n = \frac{1}{1-\sum_{n=1}^{\infty} f_{jj}^{(n)} S^n}$$

Putting $S = 1$ we have,

$$\sum_{n=0}^{\infty} p_{jj}^{(n)} = \frac{1}{1-\sum_{n=1}^{\infty} f_{jj}^{(n)}}$$

For persistency, $f_j = \sum_{n=1}^{\infty} f_{jj}^{(n)} = 1$

$$\sum_{n=0}^{\infty} p_{jj}^{(n)} = \frac{1}{1-1} = \frac{1}{0} = \infty$$

For transiency, $f_j = \sum_{n=1}^{\infty} f_{jj}^{(n)} \lhd 1$

$$\sum_{n=0}^{\infty} p_{jj}^{(n)} = \frac{1}{1-(anumberlessthan1)} \lhd \infty$$

Multiplying $P(S) = \frac{1}{1-F(S)}$ by $1-S$ and then taking the limit as $S \to 1$

$$\lim_{S \to 1}(1-S) P(S) = \lim_{S \to 1} \frac{1-S}{1-F(S)}$$

$$= \lim_{S \to 1} \frac{1-S}{1-F(S)} = \frac{1-1}{1-F(S)} = \frac{1-1}{1-1} = \frac{0}{0}$$ which is undetermined

Using the lôhopital rule;

$$\lim_{S \to 1} \frac{1-S}{1-F(S)} = \lim_{S \to 1} \frac{\frac{d}{dS}(1-S)}{\frac{d}{dS}[1-F(S)]}$$

$$= \frac{-1}{-F'(S)} = \frac{1}{F'(1)}$$

$$= \frac{1}{\mu_j}$$

For persistent state

The left hand side:

$$= \lim_{S \to 1} (1-S)P(S)$$

$$= \lim_{S \to 1} (1-S)\{p_{jj}^{(0)} + p_{jj}^{(1)}S + p_{jj}^{(2)}S^2 + \cdots + p_{jj}^{(n-1)}S^{(n-1)} + p_{jj}^{(n)}S^n + \cdots$$

$$-p_{jj}^{(0)}S - p_{jj}^{(1)}S^2 - \cdots - p_{jj}^{(n-1)}S^{(n)} + p_{jj}^{(n)}S^{n+1}\}$$

$$= \lim_{n \to \infty} p_{jj}^{(n)}$$

$$\lim_{n \to \infty} p_{jj}^{(n)} = \frac{1}{\mu_j}$$

For persistent state E_j ;

$$f_j = \sum_{n=1}^{\infty} f_{jj}^{(n)} = 1; \sum_{n=0}^{\infty} p_{jj}^{(n)} = \infty$$

For non-null, $\mu_j \lhd \infty \therefore \lim_{n \to \infty} p_{jj}^{(n)} = \frac{1}{\mu_j}$

For null,

$$\mu_j \lhd \infty \therefore \lim_{n \to \infty} p_{jj}^{(n)} = \frac{1}{\mu_j}$$

For transient state E_j

$$f_j = \sum_{n=1}^{\infty} f_{jj}^{(n)} \lhd 1; \sum_{n=0}^{\infty} p_{jj}^{(n)} \lhd \infty$$

$$\lim_{S \to 1} (1-S)P(S) = \lim_{S \to 1} \frac{(1-S)}{1-F(S)}$$

$$\lim_{S \to 1} p_{jj}^{(n)} = \frac{1-1}{1-F(1)} = 0$$

Summary of the limit theorem

$$\lim_{n \to \infty} p_{jj}^{(n)} = \begin{cases} 0 \text{ if } E_j \text{ is transient and/or null persistent} \\ 1/\varepsilon_j \text{ if } E_j \text{ is non-null persistent} \end{cases}$$

Periodicity – a state E_j is of period d when $d = \gcd\{n : p_{jj}^{(n)} \rhd 0\}$

When $d = 1$, then E_j is said to be aperiodic.

Ergodic – a state E_j is ergodic if it is persistent, non-null and aperiodic.

2.2.6 Classification of the Markov chains

Definitions

Reachability – a state E_k can be reached from state E_j if there exists a positive integer such that, the probability of moving from state E_j to E_k is greater than zero.

$p_{jk}^{(n)} \rhd 0$

Communicating states – two states E_j and E_k are said to be communicating if E_j can be reached E_k and vice versa.

Theorem: if E_k can be reached from E_j, and E_j can be reached from E_i, then E_k can be reached from E_i.

A set C of states is closed if no state outside it can be reached from set C or in any state in set C.

- **Absorbing Markov chains**

  A Markov chain is said to be absorbing if the chain has at least one absorbing state.

  A state E_j is absorbing if $p_{jj} = 1$ and $p_{kj} = 0$ for $k \neq j$

- **Irreducible Markov chain**

  A Markov chain is irreducible if there exists no closed set other than itself.

  A Markov chain is irreducible if every state can be reached from every other state.

Definition of same types

Two states are of the same type if:

1. Both are persistent or both are transient.
2. If persistent, both are either null or both are non-null.
3. Both are of the same period.

Theorem:- All states in an irreducible Markov chain are of the same type

2.2.7 Invariant (stationary) distribution

A probability distribution $\{\pi_k : k = 1, 2, \cdots\}$ is stationary or invariant if:

$$\pi_k = \sum_j \pi_j p_{jk}$$

Such that;

$$\sum_k \pi_k = 1$$

Or $\pi_j = \sum_i \pi_i p_{ij}$

Or $\pi_i = \sum_h \pi_h p_{hi}$

In general,

$$\pi_k = \sum_j \pi_j p_{jk} = \sum_j \pi_j p_{jk}{}^{(n)}$$

In short form;

$$\pi = P'\pi$$

$$\pi' = (P'\pi)'$$

$$\pi' = \pi' P$$

Theorem:- if a Markov chain is irreducible and ergodic then there exists limit

$$\pi_k = \lim_{n \to \infty} p_{jk}{}^{(n)}$$

This is independent of the initial state Ej

This π_k is an invariant distribution, that is, it satisfies:

$$\pi_k = \sum_j \pi_j p_{jk}$$

$$\sum \pi_k = 1$$

$$\pi_k \triangleright 0$$

CHAPTER III

NO CLAIMS DISCOUNT SYSTEMS

3.1 INTRODUCTION

In this chapter, different systems have been looked at with various levels, from the 3 by 3 stage NCD system to the 10 by 10 stage NCD system.

The rules applied in each of the systems have also been analysed and systems whose rules are fair to both parties, that is, the insured and the insurer, have been created.

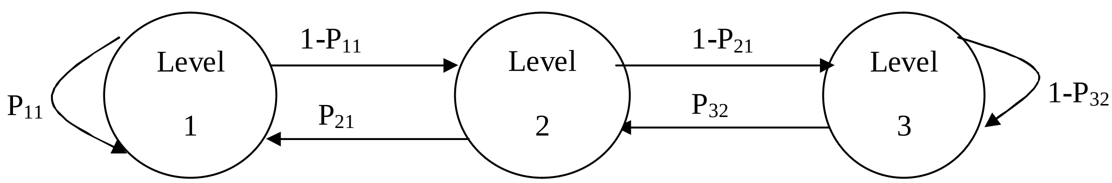
3.2 THE THREE STAGE BMS

Das and Basu (2003) consider two scenarios, explaining the movement of drivers between various levels of discount in a three stage NCD system:

System 1

The movement is by one level. Drivers move up one level to a higher discount level when they make no claim, unless they are already at the highest level, and in that case they stay there. In case of claim(s), a driver moves down by one level, unless already at the lowest level where he stays.

Transition graph 3.2.0



Transition matrix 3.2.0

$$\begin{pmatrix} P_{11} & 1-P_{11} & 0 \\ P_{21} & 0 & 1-P_{21} \\ 0 & P_{32} & 1-P_{32} \end{pmatrix}$$

$$\pi_K = \sum_j \pi_j P_{jk}$$

$$\sum \pi_k = 1$$

$$\pi_k \rhd 0$$

$$\pi' = \pi' P \text{ or } \pi = P' \pi$$

Using $\pi = P' \pi$ we have;

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} = \begin{pmatrix} P_{11} & 1-P_{11} & 0 \\ P_{21} & 0 & 1-P_{21} \\ 0 & P_{32} & 1-P_{33} \end{pmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}$$

$$(i)\pi_1 = \pi_1 P_{11} + \pi_2 P_{12}$$
$$(ii)\pi_2 = \pi_1(1-P_{11}) + \pi_3 P_{32}$$
$$(iii)\pi_3 = \pi_2(1-P_{21}) + \pi_3(1-P_{32})$$
$$(iv)\pi_1 + \pi_2 + \pi_3 = 1$$

From (iii), we have:

$$\pi_3 - \pi_3(1-P_{32}) = \pi_2(1-P_{21})$$
$$\pi_3[1-(1-P_{32})] = \pi_2(1-P_{21})$$
$$\pi_3 P_{32} = \pi_2(1-P_{21})$$
$$\pi 3 = \frac{\pi_2(1-P_{21})}{P_{32}}$$

Replacing π_3 value / equation into (ii), we have:

$$\pi_2 = \pi_1(1-P_{11}) + \frac{\pi_2(1-P_{21})}{P_{32}} P_{32}$$
$$\pi_2 = \pi_1(1-P_{11}) + \pi_2(1-P_{21})$$
$$\pi_1(1-P_{11}) = \pi_2 - \pi_2(1-P_{21})$$
$$\pi_1(1-P11) = \pi 2[1-(1-P21)]$$
$$\pi_1(1-P_{11}) = \pi_2(1-1+P_{21})$$
$$\pi_1(1-P_{11}) = \pi_2 P_{21}$$
$$\pi_1 = \frac{\pi_2 P_{21}}{(1-P_{11})}$$

Replacing the value for π_1 and π_3 into the equation (iv), we have;

$$\frac{\pi_2 P_{21}}{1-P_{11}} + \pi_2 + \frac{\pi_2(1-P_{21})}{P_{32}} = 1$$
$$\pi_2 \left[\frac{P_{21}}{1-P_{11}} + 1 + \frac{(1-P_{21})}{P_{32}} \right] = 1$$

$$\pi_2 = \left[\frac{P_{21}P_{32} + P_{32}(1-P_{11}) + (1-P_{21})(1-P_{11})}{(1-P_{11})P_{32}} \right]$$

$$\pi_2 = \frac{(1-P_{11})P_{32}}{P_{21}P_{32} + P_{32}(1-P_{11}) + (1-P_{21})(1-P_{11})}$$

The denominator;

$$= P_{21}P_{32} + P_{32}(1-P_{11}) + (1-P_{21})(1-P_{11})$$
$$= P_{21}P_{32} + P_{32} - P_{32}P_{11} + 1 - P_{11} - P_{21} + P_{21}P_{11}$$
$$= P_{21}P_{32} - P_{21} + P_{21}P_{11} + P_{32} - P_{32}P_{11} + 1 - P_{11}$$
$$= P_{21}(P_{32} - 1 + P_{11}) + (P_{32} + 1)(1 - P_{11})$$
$$= -P_{21}(1 - P_{11} - P_{32}) + (1 - P_{11})(P_{32} + 1)$$
$$= (1 - P_{11})(1 + P_{32}) - P_{21}(1 - P_{11} - P_{32})$$

Therefore, $\pi_2 = \dfrac{P_{32}(1-P_{11})}{(1-P_{11})(1+P_{32}) - P_{21}(1-P_{11}-P_{32})}$

Solution for π_1 ;

$$\pi_1 = \frac{\pi_2 P_{21}}{(1-P_{11})} = \frac{P_{32}(1-P_{11})}{(1-P_{11})(1+P_{32}) - P_{21}(1-P_{11}-P_{32})} \times \frac{P_{21}}{(1-P_{11})}$$

$$\pi_1 = \frac{P_{32}P_{21}}{(1-P_{11})(1+P_{32}) - P_{21}(1-P_{11}-P_{32})}$$

Solution for π_2;

$$\pi_2 = \frac{(1-P_{11})}{P_{21}}\pi_1$$

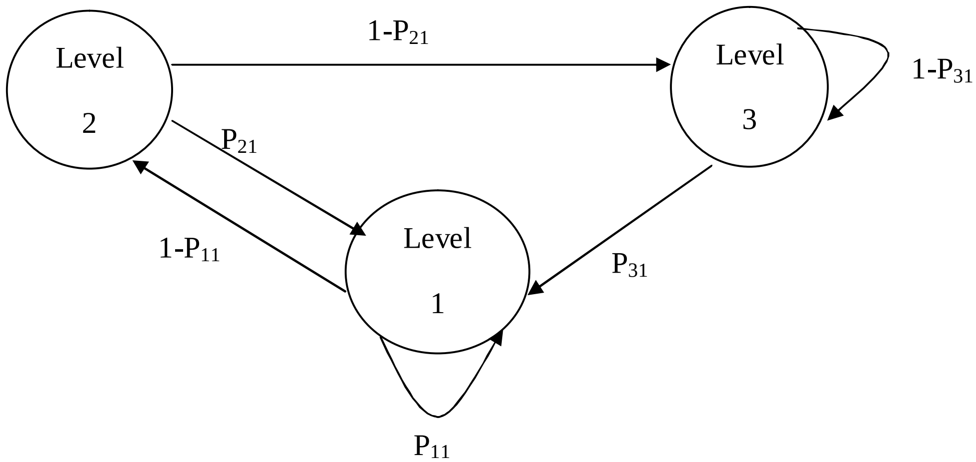
Solution for π_3;

$$\pi_3 = 1 - \pi_1 - \pi_2$$

System 2

The second scenario that Das and Basu (2003) look at is whereby, in case of no claim, the movement is the same as that of system 1 and in case of claim(s), the policyholder moves down to the lowest level of claim, that is level 1, no matter what level he is at.

Transition graph 3.2.1



Transition matrix 3.2.1

$$\begin{pmatrix} P_{11} & 1-P_{11} & 0 \\ P_{21} & 0 & P_{21} \\ P_{31} & 0 & 1-P_{31} \end{pmatrix}$$

Stationary solution

$$\pi' = \pi' P$$

$$[\pi_1 \pi_2 \pi_3] = [\pi_1 \pi_2 \pi_3] \begin{pmatrix} P_{11} & 1-P_{11} & 0 \\ P_{21} & 0 & P_{21} \\ P_{31} & 0 & 1-P_{31} \end{pmatrix}$$

$$i)\pi_1 = \pi_1 P_{11} + \pi_2 P_{21} + \pi_3 P_{31}$$
$$ii)\pi_2 = \pi_1 (1-P_{11})$$
$$iii)\pi_3 = \pi_2 (1-P_{21}) + \pi_3 (1-P_{31})$$
$$iv)\pi_1 + \pi_2 + \pi_3 = 1$$

$$From(iii)$$
$$\pi_3 - \pi_3 (1-P_{31}) = \pi_2 (1-P_{21})$$
$$\pi_3 [1-(1-P_{31})] = \pi_2 (1-P_{21})$$
$$\pi_3 P_{31} = \pi_2 (1-P_{21})$$
$$\pi_3 = \frac{\pi_2 (1-P_{21})}{P_{31}}$$

$$From(ii), \pi_1 = \frac{\pi_2}{1-P_{11}}$$

Replacing the values/equations for π_1 and π_3 into equation (iv);

$$\frac{\pi_2}{1-P_{11}} + \pi_2 + \pi_2 \frac{(1-P_{21})}{P_{31}} = 1$$

$$\pi_2 \left[\frac{1}{1-P_{11}} + 1 + \frac{1-P_{21}}{P_{31}} \right] = 1$$

$$\pi_2 = \frac{P_{31}(1-P_{11})}{P_{31} + P_{31}(1-P_{11}) + (1-P_{11})(1-P_{21})}$$

Denominator

$$= P_{31} + P_{31}(1-P_{11}) + (1-P_{11})(1-P_{21})$$
$$= P_{31} + P_{31} - P_{31}P_{11} + 1 - P_{21} - P_{11} + P_{11}P_{21}$$
$$= P_{31} + (1-P_{11})(1-P_{21} + P_{31})$$

Therefore;

$$\pi_1 = \frac{\pi_2}{1-P_{11}} = \frac{P_{31}(1-P_{11})}{P_{31} + P_{31}(1-P_{11}) + (1-P_{11})(1-P_{21})} \times \frac{1}{1-P_{11}}$$

$$\pi_1 = \frac{P_{31}}{P_{31} + P_{31}(1-P_{11}) + (1-P_{11})(1-P_{21})}$$

$$\pi_2 = (1-P_{11})\pi_1$$

$$\pi_3 = 1 - \pi_1 - \pi_2$$

Then Das and Basu (2003) describe through an example a NCD scheme defined by system 1. From the results and data obtained they discuss the features of the scheme and variations that can make it better. They discover that the set up is highly biased against good drivers, meaning the scheme is not ideal / suitable. The possible variations to improve the scheme according to them include;

- Making the variation in the probability of accident between good and bad drivers big enough.
- Severity of accidents or magnitude of claims should be brought into consideration.
- There should be higher number of categories and magnitudes of discounts and number of groups or categories of drivers.
- A driver causing an accident would not claim unless he believes that payback for loss (reimbursement) would offset (cancel/reduce the effect) the possible loss of NCD.

System 3.

This system has been used as a question by the Institute and Faculty of Actuaries Examination. CT 4 ï September 27th,2012, question 5:

A NCD system operates with three levels of discount, 0%, 15% and 40%. If a policyholder makes no claim during the year he moves up a level of discount (or remains at the maximum level). If he makes one claim during the year he moves down one level of discount (or remains at the minimum level), and if he makes two or more claims he moves down to or remains at the minimum level.

The probability for each policyholder of making two or more claims in a year is 25% of the probability of making only one claim.

The long term probability of being at the 15% level is the same as the long term probability of being at the 40% level.

The probability of one claim $= C$

The probability of two or more claims $= \beta C$

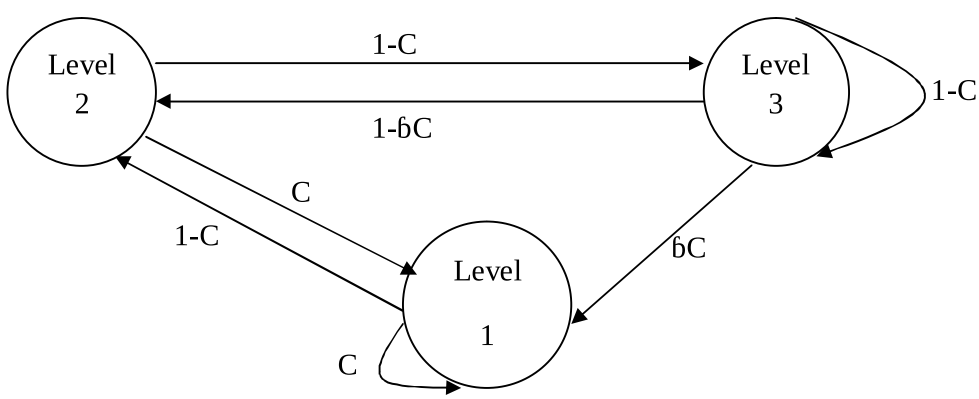
The probability of no claim $= 1 - C$

Level 1 = 0% discount

Level 2 = 15% discount

Level 3 = 40% discount

Transition graph 3.2.2



Transition matrix 3.2.2

$$\begin{pmatrix} C & 1-C & 0 \\ C & 0 & 1-C \\ \beta C & (1-\beta C) & 1-C \end{pmatrix}$$

Where $\beta = \dfrac{1}{5}$

System 3 is a special case, as it takes into consideration the frequency of the claims made by policyholders.

Remark

The system 1 and 2 illustrated by Das and Basu (2003) is a set up that is highly biased against good drivers, thus, the NCD is not ideal.

The system illustrated in the Institute and Faculty of Actuaries Examination, CT4 - September 27th, 2012, gives better and more options to drivers as opposed to system 1 or 2 solely.

For improvement of the scheme, according to Das and Basu, a possible variation would be an increase in the number of categories and magnitudes of discounts and the number of groups or categories of drivers. Another possible variation is to bring into consideration the severity of accidents or magnitudes of claims.

An example of an ideal 3 by 3 stage NCD system would be:

Level	Level occupied if;		
	$n=0$	$n=1$	$n \geq 2$
3	3	2	1
2	3	1	1
1	2	1	1

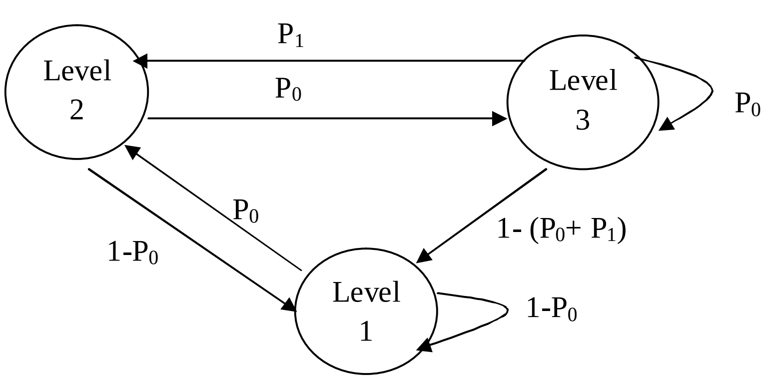
Let;

The probability of no claim = P_0

The probability of one claim = P_1

The probability of two or more claims $=1-(P_0 - P_1)$

<u>Transition graph 3.2.3</u>



Transition matrix 3.2.3

$$\begin{bmatrix} 1-P_0 & P_0 & 0 \\ 1-P_0 & 0 & P_0 \\ 1-(P_0+P_1) & P_1 & P_0 \end{bmatrix}$$

3.3.THE FOUR STAGE NCD

System 1

The Institute and Faculty of Actuaries Examination, CT 4 ï April 2009, question number 12 ï A motor insurer operates a NCD system with the following levels of discount, 0%, 25%, 50%, 60%. The rule governing a policyholderȏs discount level based upon the number of claims made in the previous year is as follows:

- Following a year with no claims, the policyholder moves up one level of discount or remains at the 60% level.
- Following a year with one claim, the policyholder moves down one level of discount or remains at the 0% level.
- Following a year with two or more claims, the policyholder moves down two discount levels (subject to a limit of the 0% discount level).

The number of claims made by a policyholder in a year is assumed to follow a Poisson distribution with mean 0.30

Probability of no claims $= \frac{e^{-\lambda}\lambda^n}{n!}$

Since n = 0, probability of no claims $= \frac{e^{-\lambda}\lambda^0}{0!} = e^{-\lambda}$

P (0 claims) $= e^{-0.3} = 0.740818$

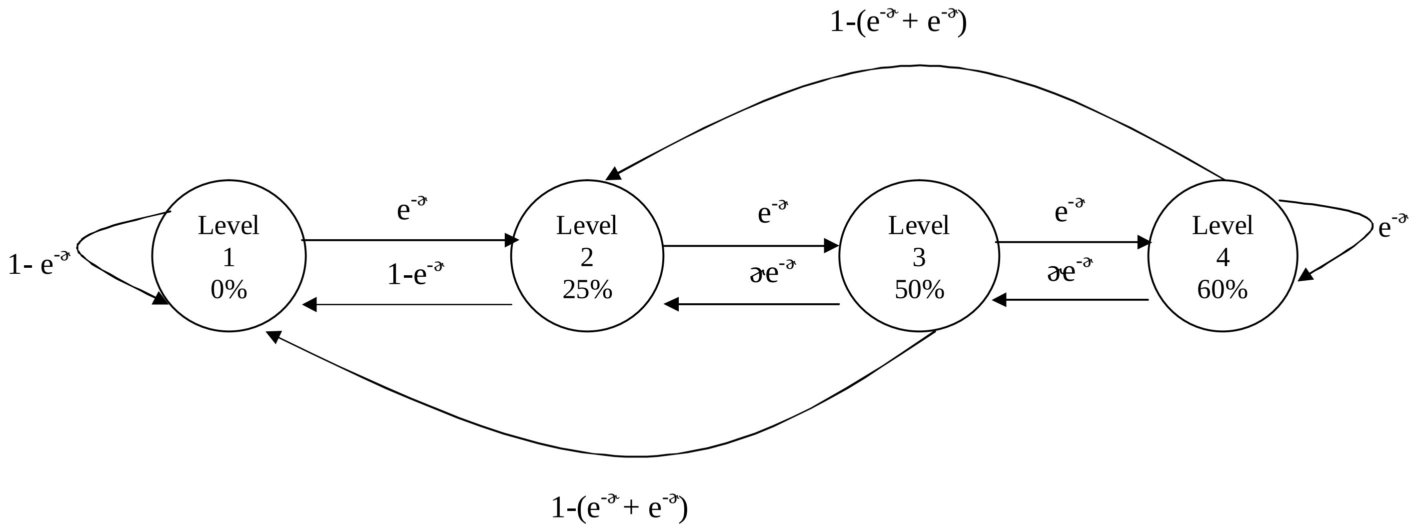
Probability of one claim $= \frac{e^{-\lambda}\lambda^1}{1!} = \lambda e^{-\lambda}$

P (1 claim) $= 0.3e^{-0.3} = 0.222245$

Probability of two or more claims = 1- 0.740818-0.222245

P (2 or more claims) =0.0369936

Transition graph 3.3.0



Transition matrix 3.3.0

Levels	1	2	3	4
1	$\lambda e^{-\lambda} + 1 - (e^{-\lambda} + \lambda e^{-\lambda})$	$e^{-\lambda}$	0	0
2	$\lambda e^{-\lambda} + 1 - (e^{-\lambda} + \lambda e^{-\lambda})$	0	$e^{-\lambda}$	0
3	$1 - (e^{-\lambda} + \lambda e^{-\lambda})$	$\lambda e^{-\lambda}$	0	$e^{-\lambda}$
4	0	$1 - (e^{-\lambda} + \lambda e^{-\lambda})$	$\lambda e^{-\lambda}$	$e^{-\lambda}$

when $\lambda = -0.3$ then we have the following transition matrix:

$$\begin{bmatrix} 0.26 & 0.74 & 0 & 0 \\ 0.26 & 0 & 0.74 & 0 \\ 0.04 & 0.22 & 0 & 0.74 \\ 0 & 0.04 & 0.22 & 0.74 \end{bmatrix}$$

Stationary distribution

Using $\pi = P'\pi$

$$\begin{bmatrix}\pi_1\\ \pi_2\\ \pi_3\\ \pi_4\end{bmatrix} = \begin{bmatrix}0.26 & 0.74 & 0 & 0\\ 0.26 & 0 & 0.74 & 0\\ 0.04 & 0.22 & 0 & 0.74\\ 0 & 0.04 & 0.22 & 0.74\end{bmatrix} \times \begin{bmatrix}\pi_1\\ \pi_2\\ \pi_3\\ \pi_4\end{bmatrix}$$

$i)\pi_1 = 0.26\pi_1 + 0.26\pi_2 + 0.04\pi_3$
$ii)\pi_2 = 0.74\pi_1 + 0.22\pi_3 + 0.04\pi_4$
$iii)\pi_3 = 0.74\pi_2 + 0.22\pi_4$
$iv)\pi_4 = 0.74\pi_3 + 0.74\pi_4$
$v)\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$

From equation (iv), we have;

$\pi_4 - 0.74\pi_4 = 0.74\pi_3$
$0.26\pi_4 = 0.74\pi_3$
$\pi_4 = \frac{0.74}{0.26}\pi_3 = 2.8462\pi_3$
$\pi_4 = 2.85\pi_3$

Replacing the value of π_4 into equation (iii), we have;

$\pi_3 = 0.74\pi_2 + (0.22 \times 2.85\pi_3)$
$\pi_3 = 0.74\pi_2 + 0.6262\pi_3$
$\pi_3 - 0.6262\pi_3 = 0.74\pi_2$
$0.3738\pi_3 = 0.74\pi_2$
$\pi_2 = \frac{0.3738}{0.74}\pi_3$
$\pi_2 = 0.5052\pi_3$
$\pi_2 = 0.51\pi_3$

Replacing the value of π_2 into equation (i), we have;

$\pi_1 = 0.26\pi_1 + (0.26 \times 0.51\pi_3) + 0.04\pi_3$
$\pi_1 - 0.26\pi_1 = 0.13\pi_3 + 0.04\pi_3$
$0.74\pi_1 = 0.17\pi_3$
$\pi_1 = \frac{0.17}{0.74}\pi_3$
$\pi_1 = 0.23\pi_3$

Replacing the values for $\pi_{1,}\pi_2 and\, \pi_4$ into equation (v), we have;

$$0.23\pi_3 + 0.51\pi_3 + \pi_3 + 2.85\pi_3 = 1$$
$$4.59\pi_3 = 1$$
$$\pi_3 = 0.22$$
$$\pi_1 = (0.23 \times 0.22) = 0.0501$$
$$\pi_2 = (0.51 \times 0.22) = 0.1111$$
$$\pi_4 = (2.85 \times 0.22) = 0.6210$$

Average discount $= 60\%\pi_4 + 50\%\pi_3 + 25\%\pi_2 + 100\%\pi_1$

Remark

The transition rule for the four stage NCD system above, from the Institute and Faculty of Actuaries Examination, does not take into consideration the claim frequency.

An example of an ideal 4 by 4 stage NCD system would be:

Level	Level occupied if;			
	$n = 0$	$n = 1$	$n = 2$	$n \geq 3$
4	4	3	2	1
3	4	2	1	1
2	3	1	1	1
1	2	1	1	1

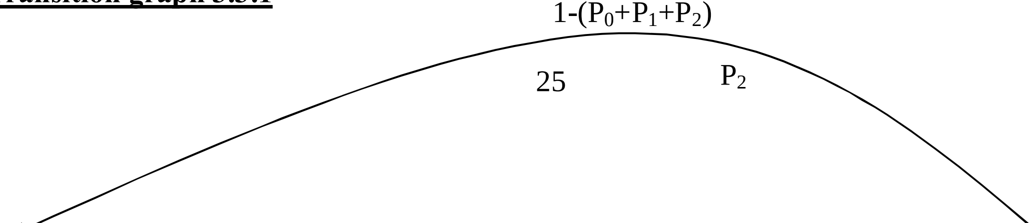
The probability of no claim $= P_0$

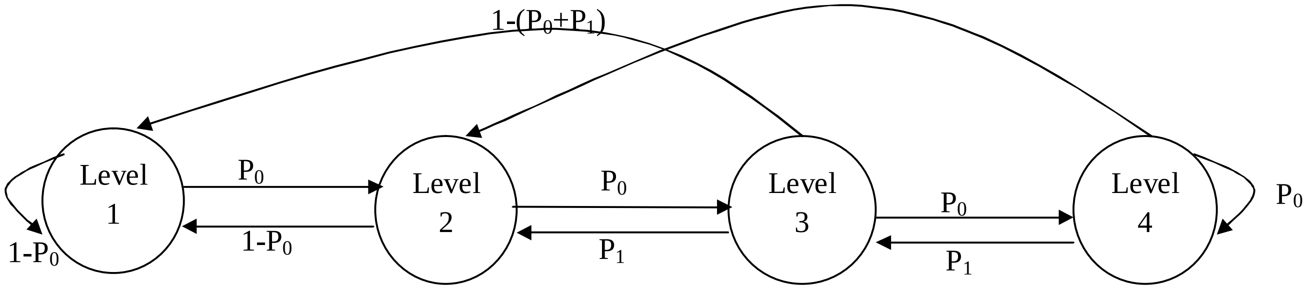
The probability of one claim $= P_1$

The probability of two claims $= P_2$

The probability of three or more claims $= 1 - (P_0 + P_1 + P_2)$

<u>Transition graph 3.3.1</u>





Transition matrix 3.3.1

$$\begin{bmatrix} 1-P_0 & P_0 & 0 & 0 \\ 1-P_0 & 0 & P_0 & 0 \\ 1-(P_0+P_1) & P_1 & 0 & P_0 \\ 1-(P_0+P_1+P_2) & P_2 & P_1 & P_0 \end{bmatrix}$$

3.4 THE FIVE STAGE BMS

System 1

Lemaire and Zi (1994) simulate and compare the automobile third party insurance merit rating systems of 22 countries, to define an óindex of toughnessô for all the systems. One of the countries whose BMS was analysed was Spain with 5 classes and premium levels, 70%, 80%, 90%, 100% and 100%. The starting level is 100%, and in case of no claim a policyholder moves up by one level. For each claim made by a policyholder, he moves all the way to the 100% level and loses all discounts.

Transition graph 3.4.0

Assumption is that the number of at-fault claims for a given policyholder conforms to a Poisson distribution with parameter λ .

The probability of no claim will be $(n=0)$;

$$= \frac{e^{-\lambda}\lambda^{n}}{n!}$$

$$= \frac{e^{-\lambda}\lambda^{0}}{0!} = e^{-\lambda}$$

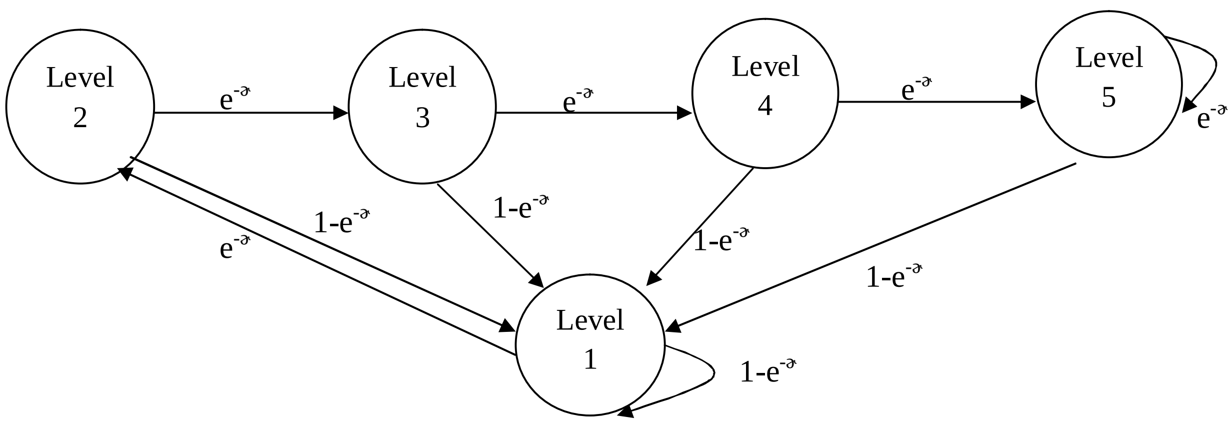
The probability of a claim, $n = 1 = (1 - e^{-\lambda})$

Level 100% means the driver pays the full premium, getting 0% discount.

Level 90% means the driver pays the full premium, getting 10% discount.

Level 80% means the driver pays the full premium, getting 20% discount.

Level 70% means the driver pays the full premium, getting 30% discount.



Transition matrix 3.4.0

$$\begin{bmatrix} 1-e^{-\lambda} & e^{-\lambda} & 0 & 0 & 0 \\ 1-e^{-\lambda} & 0 & e^{-\lambda} & 0 & 0 \\ 1-e^{-\lambda} & 0 & 0 & e^{-\lambda} & 0 \\ 1-e^{-\lambda} & 0 & 0 & 0 & e^{-\lambda} \\ 1-e^{-\lambda} & 0 & 0 & 0 & e^{-\lambda} \end{bmatrix}$$

The use of the Spain system has been discontinued by most insurers as complete rating freedom now exists.

System 2

Emamverdi et al (2013), focus on techniques / practical method for construction of optimal BMSs with reasonable penalties that can be commercially implemented.

They look at the Iranian BMS that rewards policyholders with no claim in a year by going down one level in the scale, whereas send them to the first level, the level with the largest premium as penalty in case of a claim.

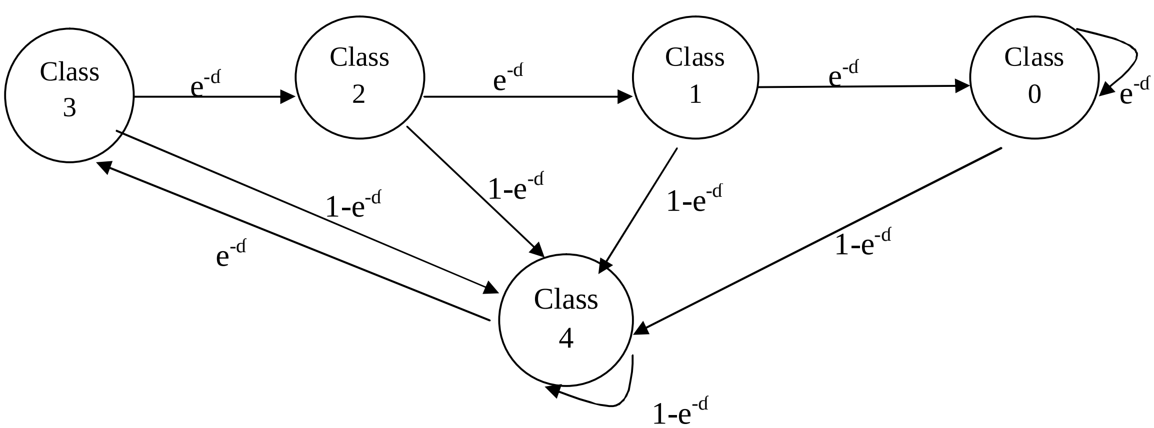
The rule of the Iranian system is shown in the following table:

Class	Class after 0 claims	Class after one or more claims
4	3	4
3	2	4
2	1	4
1	0	4
0	0	4

Class 4 is the lowest level with the lowest level of discount, while class 0 is the highest level with the highest level of discount.

The probability of no claim is $e^{-\theta}$ and the probability of one or more claims is $1-e^{-\theta}$.

Transition graph 3.4.1



Transition matrix 3.4.1

The transition probability matrix will be:

$$P(\theta) = \begin{bmatrix} e^{-\theta} & 0 & 0 & 0 & 1-e^{-\theta} \\ e^{-\theta} & 0 & 0 & 0 & 1-e^{-\theta} \\ 0 & e^{-\theta} & 0 & 0 & 1-e^{-\theta} \\ 0 & 0 & e^{-\theta} & 0 & 1-e^{-\theta} \\ 0 & 0 & 0 & e^{-\theta} & 1-e^{-\theta} \end{bmatrix}$$

Remark

The Spanish system from Lemaire and Zi (1994) and the Iranian system from Emamverdi et al (2013) that we looked at, the transition rules are the same for both systems. The set up is biased against careful drivers, as a driver with only one claim is treated the same as a driver with two or more claims.

An example of a five stage BMS that takes into consideration the frequency of claims would be one with the following transition rules and with level 5 being the highest discount level and level 1 being the lowest.

Level	Level occupied if;				
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n \geq 4$
5	5	4	3	2	1
4	5	3	2	1	1
3	4	2	1	1	1
2	3	1	1	1	1
1	2	1	1	1	1

Let:

The probability of no claim $= P_0$

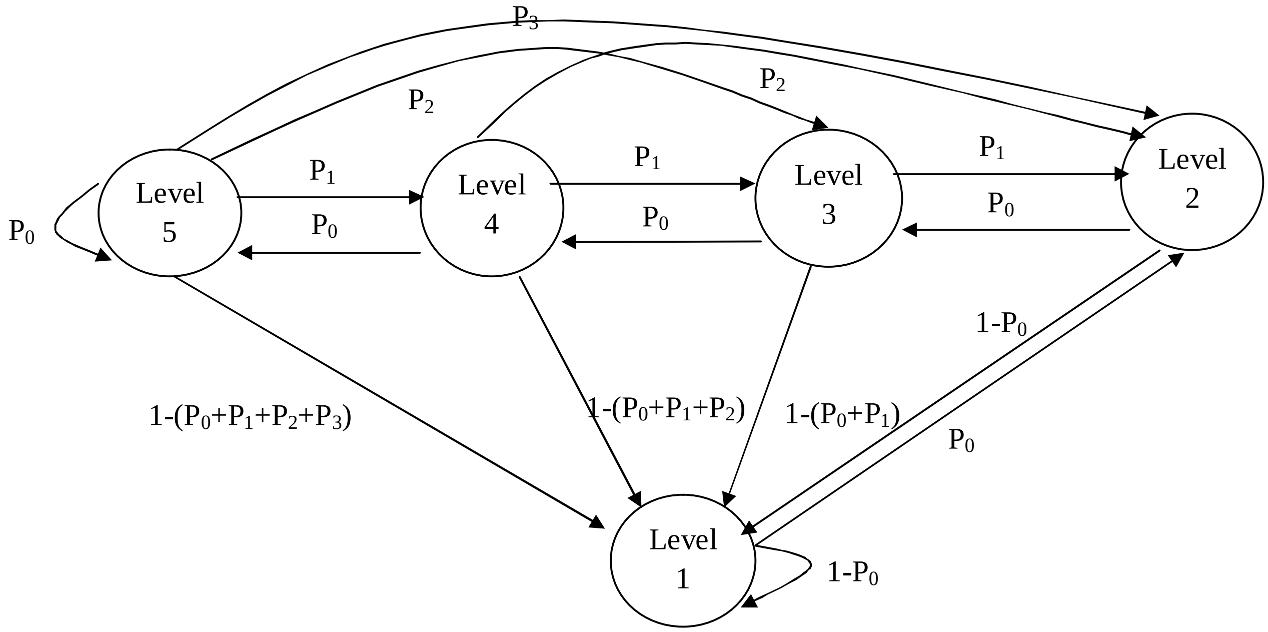
The probability of one claim $= P_1$

The probability of two claims $= P_2$

The probability of three claims $= P_3$

The probability of four or more claims $= 1-(P_0 + P_1 + P_2 + P_3)$

<u>Transition graph 3.4.2</u>



Transition matrix 3.4.2

$$\begin{bmatrix} P_0 & P_1 & P_2 & P_3 & 1-(P_0+P_1+P_2+P_3) \\ P_0 & 0 & P_1 & P_2 & 1-(P_0+P_1+P_2) \\ 0 & P_0 & 0 & P_1 & 1-(P_0+P_1) \\ 0 & 0 & P_0 & 0 & 1-P_0 \\ 0 & 0 & 0 & P_0 & 1-P_0 \end{bmatrix}$$

The BMS is suitable as it considers the claims frequency.

3.5 THE SIX STAGE BMS

System 1

Nath and Sinha (2014) try to find out the probabilities of claims by different categories of policyholders (motorists). They inspect the desirability of the multi-layer premium system of the Insurance Regulatory and Development Authority (IRDA).

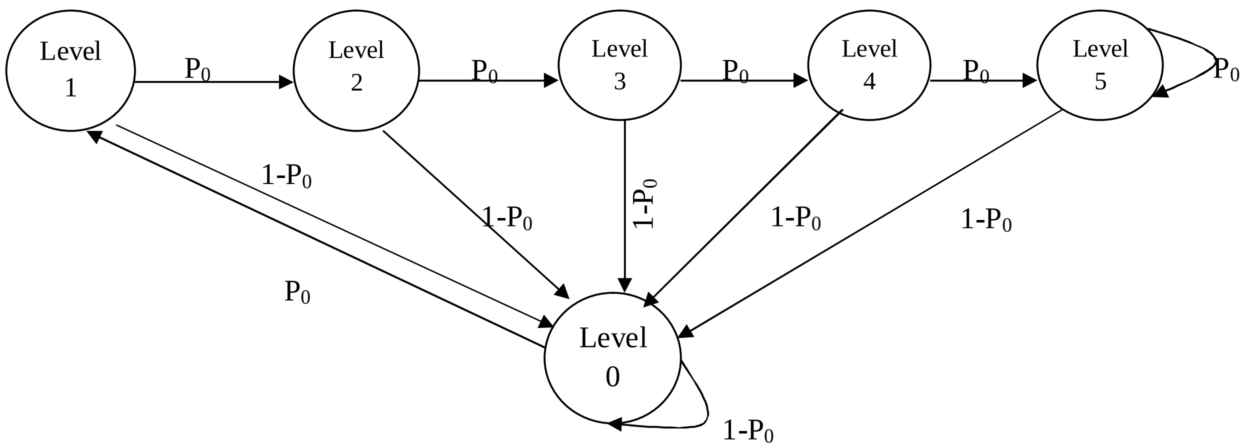
The levels of NCD system of IRDA is as follows:

Levels / Age of vehicle	No Claim Discount Saving

5	50%
4	45%
3	35%
2	25%
1	20%
0	00%

Each year at renewal, a policyholder moves up to the next level of NCD if he hasnǿt made a claim for an accident where he is at fault in that year. If a policyholder makes a claim for an accident where he is at fault he moves down to level zero unless he is on maximum NCD for life. The NCD level does not change when a policyholder makes a claim for something that is not his fault like his car or motorcycle is stolen or damaged by storm.

Transition graph 3.5.0



Probability of no claim is P_0 and the probability of at least one claim ($1-P_0$)

Transition probability matrix 3.5.0

$$\begin{bmatrix} 1-P_0 & P_0 & 0 & 0 & 0 & 0 \\ 1-P_0 & 0 & P_0 & 0 & 0 & 0 \\ 1-P_0 & 0 & 0 & P_0 & 0 & 0 \\ 1-P_0 & 0 & 0 & 0 & P_0 & 0 \\ 1-P_0 & 0 & 0 & 0 & 0 & P_0 \\ 1-P_0 & 0 & 0 & 0 & 0 & P_0 \end{bmatrix}$$

Stationary distribution

Using $\pi' = \pi' P$ we have;

$$[\pi_0 \quad \pi_1 \quad \pi_2 \quad \pi_3 \quad \pi_4 \quad \pi_5] = [\pi_0 \quad \pi_1 \quad \pi_2 \quad \pi_3 \quad \pi_4 \quad \pi_5]\begin{bmatrix} 1-P_0 & P_0 & 0 & 0 & 0 & 0 \\ 1-P_0 & 0 & P_0 & 0 & 0 & 0 \\ 1-P_0 & 0 & 0 & P_0 & 0 & 0 \\ 1-P_0 & 0 & 0 & 0 & P_0 & 0 \\ 1-P_0 & 0 & 0 & 0 & 0 & P_0 \\ 1-P_0 & 0 & 0 & 0 & 0 & P_0 \end{bmatrix}$$

$(i)\pi_0 = \pi_0(1-P_0) + \pi_1(1-P_0) + \pi_2(1-P_0) + \pi_3(1-P_0) + \pi_4(1-P_0) + \pi_5(1-P_0)$
$\pi_0 = 1-P_0(\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5)$
$(ii)\pi_1 = \pi_0 P_0$
$(iii)\pi_2 = \pi_1 P_0$
$(iv)\pi_3 = \pi_2 P_0$
$(v)\pi_4 = \pi_3 P_0$
$(vi)\pi_5 = \pi_4 P_0 + \pi_5 P_0 = P_0(\pi_4 + \pi_5)$
$(vii)\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 = 1$

Replacing the equation (vii) into (i), we have; $\pi_0 = 1-P_0$

Replacing the value of $\pi_0 = 1-P_0$ into equation (ii), we have; $\pi_1 = (1-P_0)P_0$

Replacing the value of π_1 into equation (iii), we get;

$\pi_2 = (1-P_0)P_0 \bullet P_0 = (1-P_0)P_0^2$

Replacing the value of π_2 equation (iv), we have;

$\pi_3 = (1-P_0)P_0^2 \bullet P_0 = (1-P_0)P_0^3$

Replacing the value of π_3 into equation (v), we get;

$\pi_4 = (1-P_0)P_0^3 \bullet P_0 = (1-P_0)P_0^4$

Replacing the value of π_4 into equation (vi), we have;

$$\pi_5 = [(1-P_0)P_0^4 \bullet P_0] + \pi_5 P_0$$
$$\pi_5 - \pi_5 P_0 = (1-P_0)P_0^5$$
$$\pi_5(1-P_0) = (1-P_0)P_0$$
$$\pi_5 = P_0^5$$

The average yearly premium paid, $A(P_0, m)$ in the steady state in terms of P_0 and m is;

$A(P_0, m) = m\sum_{i=0}^{5} \pi_i \times$ Percentage of discounts at different levels

$$= \frac{m}{100}(1-P_0)[100 + 80P_0 + 75P_0^2 + 65P_0^3 + 55P_0^4 + (50\frac{P_0^5}{1-P_0})$$
$$= m\left\{(1-P_0)100\% + [(1-P_0)P_0]80\% + [(1-P_0)P_0^2]75\% + [(1-P_0)P_0^3]65\% + [(1-P_0)P_0^4]55\% + P_0^5 \times 50\%\right\}$$

where m is the yearly amount of premium.

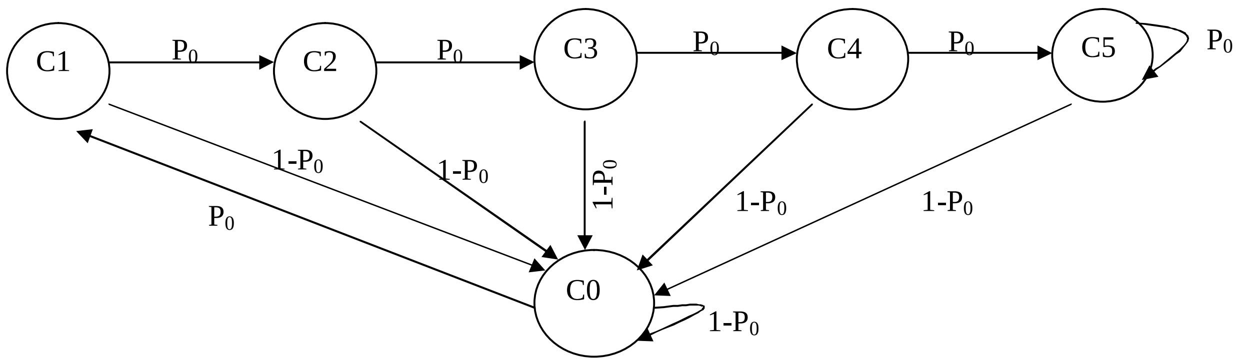
System 2

Ibiwoye et al (2011) evaluate the BMS in practice in Nigeria motor insurance industry and construct an alternative BMS that has reasonable penalties and is commercially feasible, since the Nigerian BMS is a rule of thumb as operators do not honour the industry agreed tariff.

In the Nigerian BMS, private car policyholders are grouped into six classes with premium levels 100, 80, 75, 66.7, 60 and 50 designated as C0, C1, C2, C3, C4 and C5 respectively. Reporting of a claim leads to loss of all discounts, regardless of the policyholderǿs risk class and he starts all over from class C0, where he will pay 100% of annual premium. In case of no claim, the policyholder moves to the next premium level. If a policyholder moves to another insurer at the end of the period, he keeps the same class as long as he is able to document his class with the previous insurer.

NCD class	C0	C1	C2	C3	C4	C5
Percentage Discount	0	20	25	$33\frac{1}{3}$	40	50
Percentage of premium	100	80	75	$66\frac{2}{3}$	60	50

Transition graph 3.5.1



Let $P_0 =$ Probability of no claim and $(1-P_0) =$ Probability of a claim

Transition probability matrix 3.5.1

$$\begin{bmatrix} (1-P_0) & P_0 & 0 & 0 & 0 & 0 \\ (1-P_0) & 0 & P_0 & 0 & 0 & 0 \\ (1-P_0) & 0 & 0 & P_0 & 0 & 0 \\ (1-P_0) & 0 & 0 & 0 & P_0 & 0 \\ (1-P_0) & 0 & 0 & 0 & 0 & P_0 \\ (1-P_0) & 0 & 0 & 0 & 0 & P_0 \end{bmatrix}$$

Among the 22 countries that Lemaire and Zi (1994) look at, include Hong Kong, Malaysia and Portugal.

System 3

The Portugal BMS has six number of classes with levels 70, 100, 115, 130, 145 and 200. Its starting level is 100, and for a claim free two years, a policyholder moves down a premium level and for each claim a policyholder moves a premium level. (Note ï if a policyholder does not make a claim within a year, he does not move down a premium level or up the discount level, this only happens after two consecutive claim-free years.

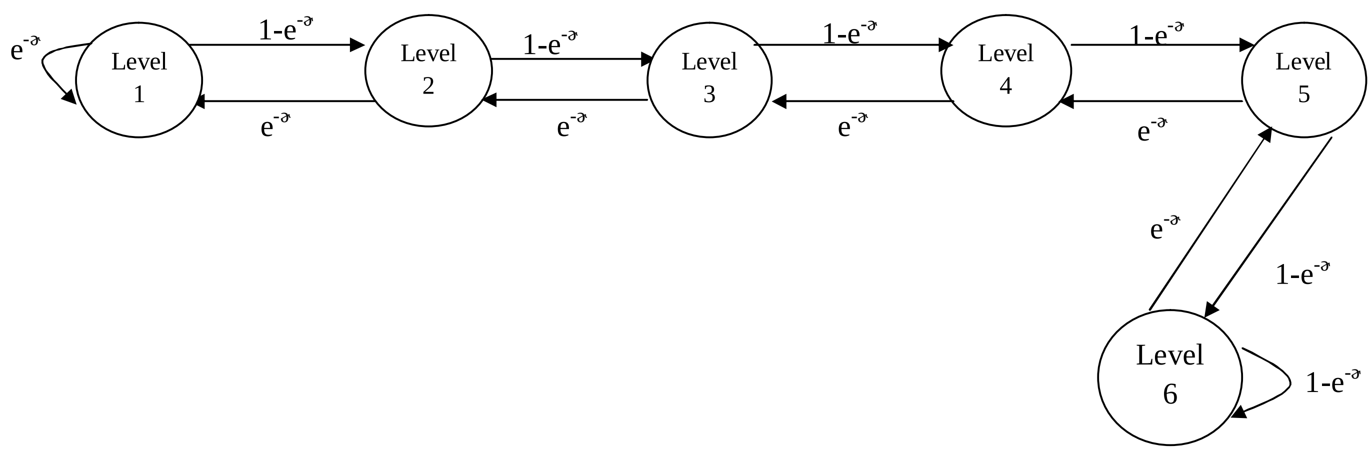
Transition graph 3.5.2

The number of at fault claims for a policyholder, conform to a Poisson distribution with parameter λ .

Probability of no claim $= e^{-\lambda}$

Probability of a claim $= 1 - e^{-\lambda}$

Level	Premium level	Discount level
1	70%	30%
2	100%	0% (starting level)
3	115%	-15%
4	130%	-30%
5	145%	-45%
6	200%	-100%



Transition probability matrix 3.5.2

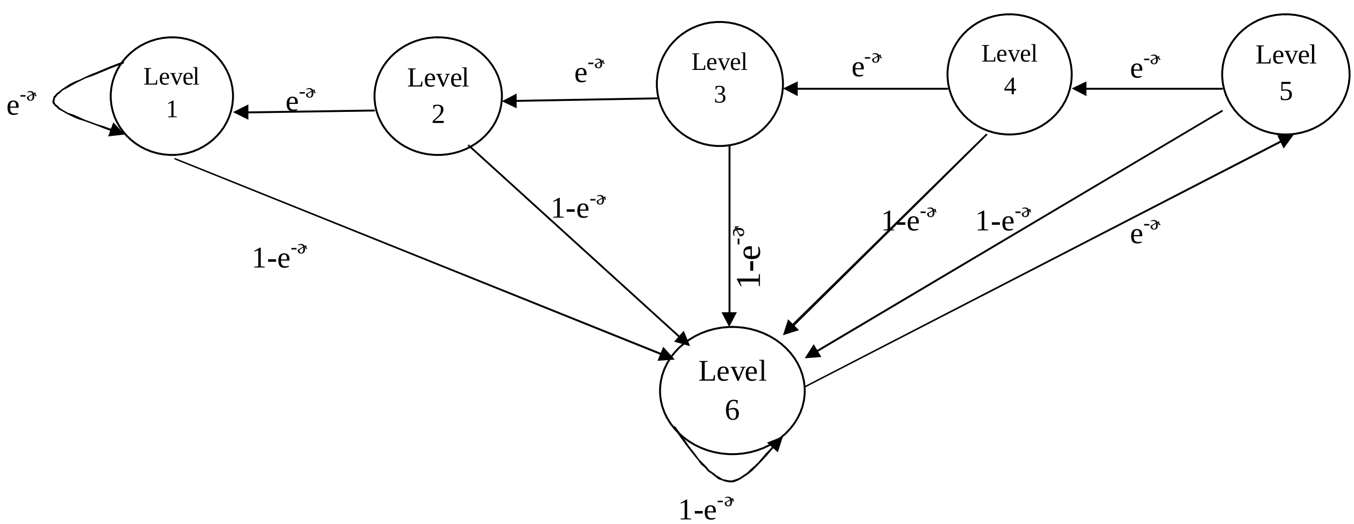
$$\begin{bmatrix} e^{-\lambda} & 1-e^{-\lambda} & 0 & 0 & 0 & 0 \\ e^{-\lambda} & 0 & 1-e^{-\lambda} & 0 & 0 & 0 \\ 0 & e^{-\lambda} & 0 & 1-e^{-\lambda} & 0 & 0 \\ 0 & 0 & e^{-\lambda} & 0 & 1-e^{-\lambda} & 0 \\ 0 & 0 & 0 & e^{-\lambda} & 0 & 1-e^{-\lambda} \\ 0 & 0 & 0 & 0 & e^{-\lambda} & 1-e^{-\lambda} \end{bmatrix}$$

System 4

The Malaysian ï Singaporeôs BMS has six numbers of classes as well with premium levels 45, 55, 61.67, 70, 75 and 100. The starting point is level 100, for a claim-free year a policyholder moves down a premium level and in case of a claim, all discounts are lost.

Level	Premium level	Discount level
1	45	55
2	55	45
3	61.67	38.33
4	70	30
5	75	25
6	100	0

Transition graph 3.5.3



Transition matrix 3.5.3

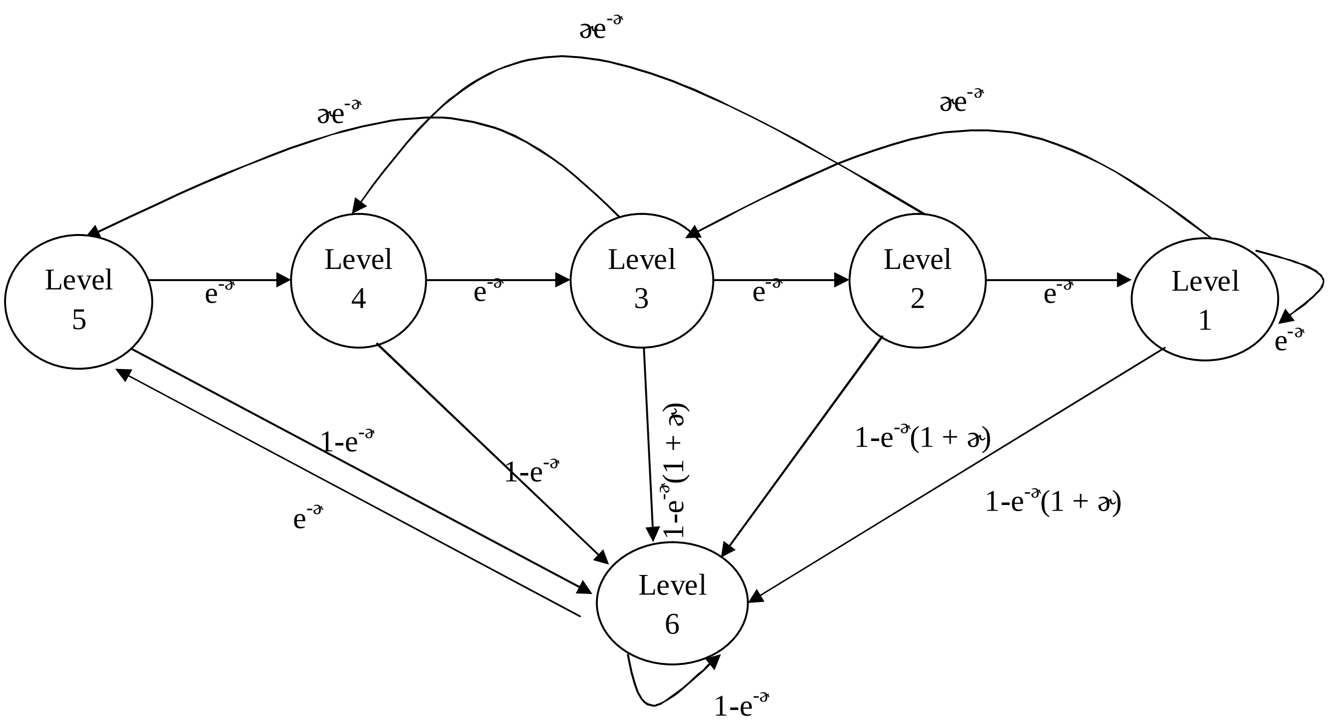
$$\begin{bmatrix} e^{-\lambda} & 0 & 0 & 0 & 0 & 1-e^{-\lambda} \\ e^{-\lambda} & 0 & 0 & 0 & 0 & 1-e^{-\lambda} \\ 0 & e^{-\lambda} & 0 & 0 & 0 & 1-e^{-\lambda} \\ 0 & 0 & e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} \\ 0 & 0 & 0 & e^{-\lambda} & 0 & 1-e^{-\lambda} \\ 0 & 0 & 0 & 0 & e^{-\lambda} & 1-e^{-\lambda} \end{bmatrix}$$

System 5

The Hong Kong BMS has six classes with levels 40, 50, 60, 70, 80 and 100 with the starting level being 100, the movement between the levels is described in the table below.

Level	Level occupied if			Premium level (percentage)	Discount level (percentage)
	$n=0$	$n=1$	$n \geq 2$		
1	1	3	6	40	60
2	1	4	6	50	50
3	2	5	6	60	40
4	3	6	6	70	30
5	4	6	6	80	20
6	5	6	6	100	0

Transition graph 3.5.4



Transition probability matrix 3.5.4

$$\begin{bmatrix} e^{-\lambda} & 0 & \lambda e^{-\lambda} & 0 & 0 & 1-e^{-\lambda}(1+\lambda) \\ e^{-\lambda} & 0 & 0 & \lambda e^{-\lambda} & 0 & 1-e^{-\lambda}(1+\lambda) \\ 0 & e^{-\lambda} & 0 & 0 & \lambda e^{-\lambda} & 1-e^{-\lambda}(1+\lambda) \\ 0 & 0 & e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} \\ 0 & 0 & 0 & e^{-\lambda} & 0 & 1-e^{-\lambda} \\ 0 & 0 & 0 & 0 & e^{-\lambda} & 1-e^{-\lambda} \end{bmatrix}$$

System 6

Asmussen (2014) in his article, modelling and performance of Bonus Malus Systems: Stationarity Versus age-correction, suggests an age-correction to the stationary distribution and present an extensive numerical study of its effects. The number of claims is assumed to conform to a Poisson distribution. He looks at the BMS in Ireland, which has six classes and the transition rules are as described below:

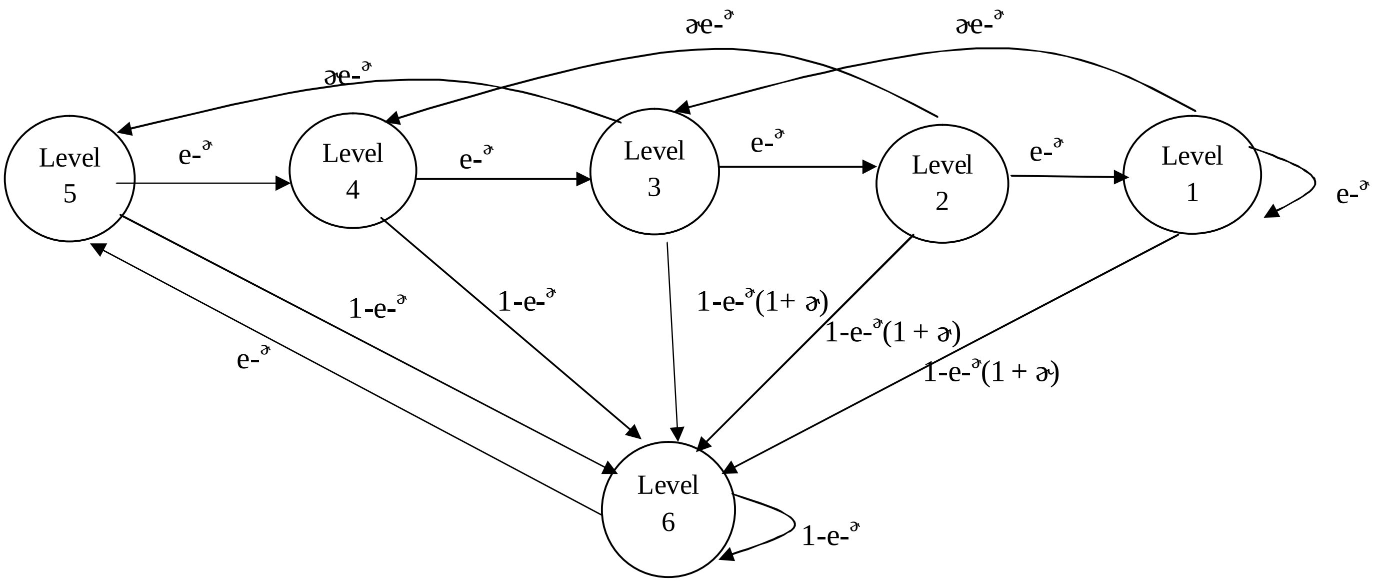
l	r_1	$n = 0$	$n = 1$	$n = 2+$
6	100	5	6	6
5	90	4	6	6
4	80	3	6	6
3	70	2	5	6
2	60	1	4	6
1	50	1	3	6

Where: $l =$ class

r_1 =premium level

n=number of claims

Transition graph 3.5.5



The probability of no claim $= e^{-\lambda}$

The probability of one claim $= \lambda e^{-\lambda}$

The probability of two or more claims $= 1-(e^{-\lambda} + \lambda e^{-\lambda})$

Transition matrix 3.5.5

$$\begin{bmatrix} 1-e^{-\lambda} & e^{-\lambda} & 0 & 0 & 0 & 0 \\ 1-e^{-\lambda} & 0 & e^{-\lambda} & 0 & 0 & 0 \\ 1-e^{-\lambda} & 0 & 0 & e^{-\lambda} & 0 & 0 \\ 1-(e^{-\lambda}+\lambda e^{-\lambda}) & \lambda e^{-\lambda} & 0 & 0 & e^{-\lambda} & 0 \\ 1-(e^{-\lambda}+\lambda e^{-\lambda}) & 0 & \lambda e^{-\lambda} & 0 & 0 & e^{-\lambda} \\ 1-(e^{-\lambda}+\lambda e^{-\lambda}) & 0 & 0 & \lambda e^{-\lambda} & 0 & e^{-\lambda} \end{bmatrix}$$

Remark

The IRDA NCD system by Nath and Sinha (2014), Nigerian BMS by Ibiwoye et al (2011) and the Malaysian ï Singapore BMS by Lemaire and Zi (1994) have 6 classes with the same transition rules explaining the movement of drivers between these classes. The systems treat drivers who make claim(s) the same, not considering the number of claims made by an individual. They bring the policyholders with claim(s) all the way back to the lowest discount level, no matter what level they are at. Such systems discourage drivers.

The Portugal BMS looked at by Lemaire and Zi (1994) has a transition looked at has a transition rule that discriminates against careful drivers. A driver with one claim in a year should not be treated the same as a driver with more than one claim.

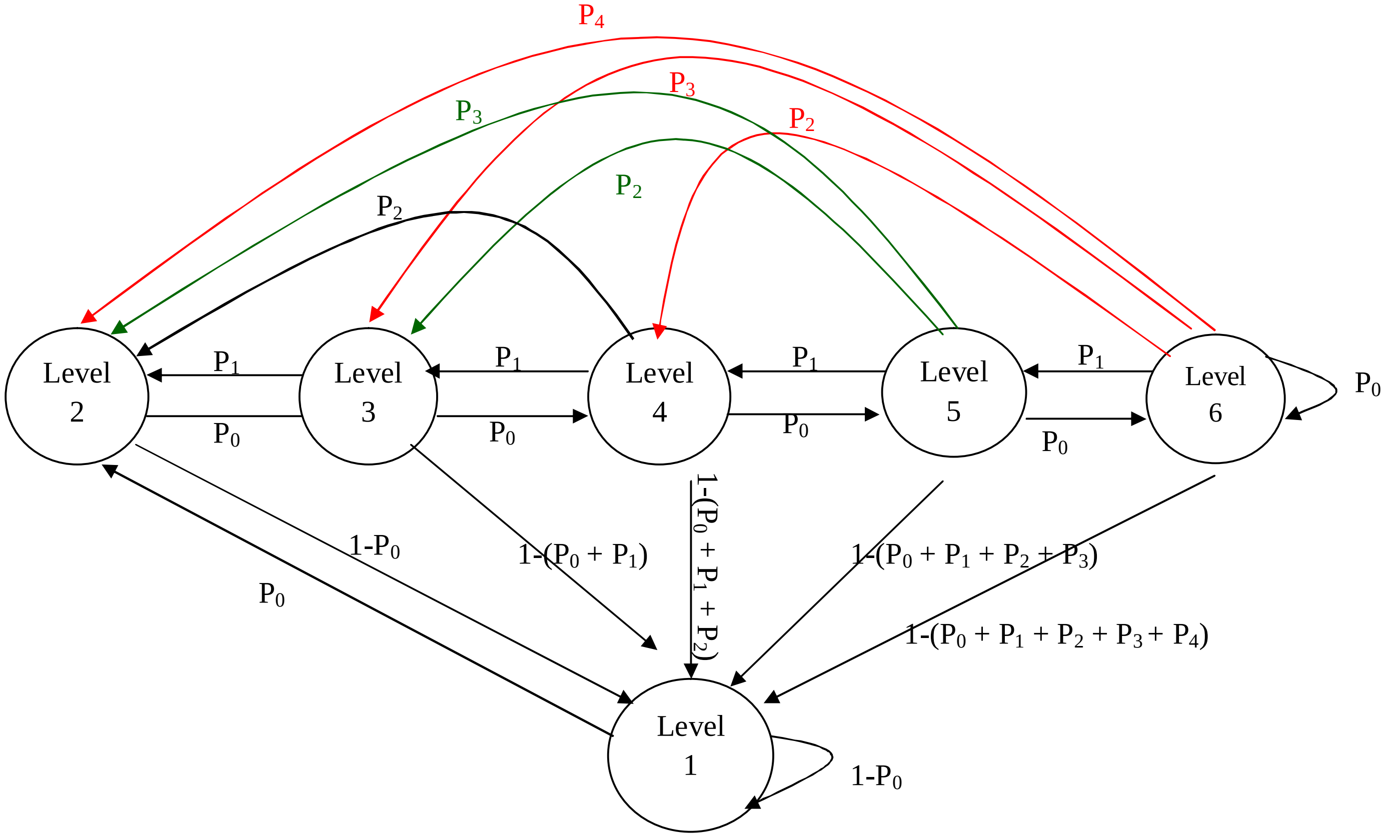
The Hong Kong BMS by Lemaire and Zi (1994) as well as the Irish system discussed in the article by Asmussen (2014), have 6 classes with similar transition rules, where a driver who makes a claim and is at level 4, 5 or 6 is treated the same as a driver who makes two or more claims at any level.

An example of an ideal six stage BMS that takes into consideration frequency of claims is as follows:

Level	Level occupied when					
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n \geq 5$
6	6	5	4	3	2	1
5	6	4	3	2	1	1
4	5	3	2	1	1	1
3	4	2	1	1	1	1
2	3	1	1	1	1	1
1	2	1	1	1	1	1

With level 6 being the highest discount level.

Transition graph 3.5.6



Transition matrix 3.5.6

$$
\begin{array}{c}6\\5\\4\\3\\2\\1\end{array}
\begin{bmatrix}
P_0 & P_1 & P_2 & P_3 & P_4 & 1-(P_0+P_1+P_2+P_3+P_4) \\
P_0 & 0 & P_1 & P_2 & P_3 & 1-(P_0+P_1+P_2+P_3) \\
0 & P_0 & 0 & P_1 & P_2 & 1-(P_0+P_1+P_2) \\
0 & 0 & P_0 & 0 & P_1 & 1-(P_0+P_1) \\
0 & 0 & 0 & P_0 & 0 & 1-P_0 \\
0 & 0 & 0 & 0 & P_0 & 1-P_0
\end{bmatrix}
$$

3.6 THE SEVEN STAGE BMS

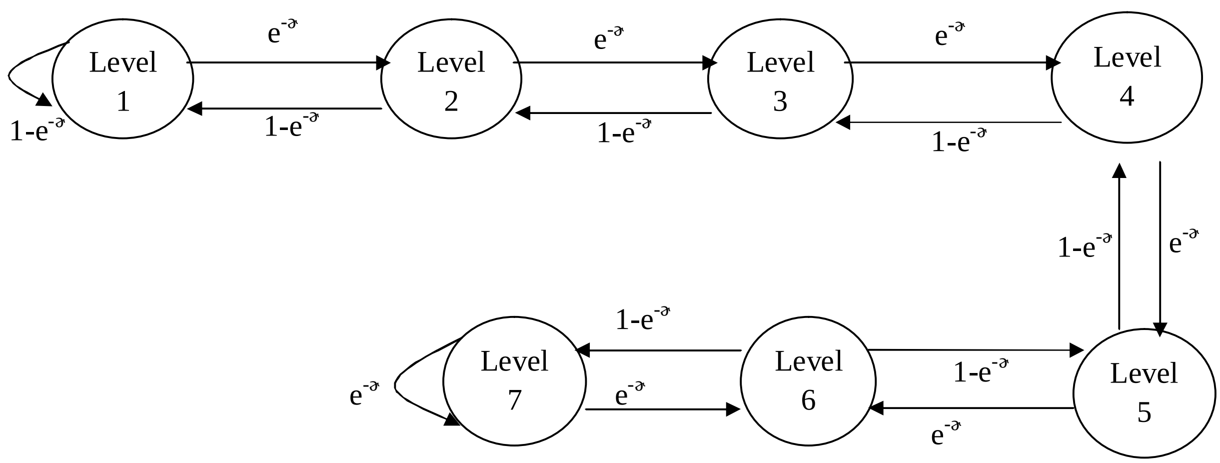
The other countries that Lemaire and Zi (1994) look at are Brazil, Kenya, Sweden, Thailand and UK.

System 1

Brazil has 7 classes with levels 65, 70, 75, 80, 90 and 100 (starting level). for a claim free year a policyholder moves down a premium level, where he pays less premium and for each claim made he moves up a premium level.

Transition graph 3.6.0

Level	Percentage of premium	Percentage of discount
1	100	0
2	90	10
3	85	15
4	80	20
5	75	25
6	70	30
7	65	35



Assumption ï the BMS has the number of at-fault claims for a given policyholder conforming to a Poisson distribution.

Probability of no claim $= e^{-\lambda}$ and the probability of a claim $= 1 - e^{-\lambda}$

Transition matrix 3.6.0

$$\begin{bmatrix} 1-e^{-\lambda} & e^{-\lambda} & 0 & 0 & 0 & 0 & 0 \\ 1-e^{-\lambda} & 0 & e^{-\lambda} & 0 & 0 & 0 & 0 \\ 0 & 1-e^{-\lambda} & 0 & e^{-\lambda} & 0 & 0 & 0 \\ 0 & 0 & 1-e^{-\lambda} & 0 & e^{-\lambda} & 0 & 0 \\ 0 & 0 & 0 & 1-e^{-\lambda} & 0 & e^{-\lambda} & 0 \\ 0 & 0 & 0 & 0 & 1-e^{-\lambda} & 0 & e^{-\lambda} \\ 0 & 0 & 0 & 0 & 0 & 1-e^{-\lambda} & e^{-\lambda} \end{bmatrix}$$

System 2

Lemaire (2013) in his article, North America Actuarial Journal looks at the Brazilian BMS again, with the same number of classes and levels. New policyholders start at level 7 (100% premium level). The transition rules are presented below:

Class	Level of premium (percentage)	Level occupied if						
		0 claims	1 claim	2 claims	3 claims	4 claims	5 claims	Ó6 claims
7	100	6	7	7	7	7	7	7
6	90	5	7	7	7	7	7	7
5	85	4	6	7	7	7	7	7
4	80	3	5	6	7	7	7	7
3	75	2	4	5	6	7	7	7
2	70	1	3	4	5	6	7	7
1	65	1	2	3	4	5	6	7

The assumption is that the distribution $(P_k; k = 0,1,2,...)$ of the number of claims of a specific driver conforms to a Poisson distribution with parameter λ , where λ is called the claim frequency of the policyholder and is assumed to be constant over time.

$$P_k = \frac{e^{-\lambda}\lambda_k}{k!}$$

The probability of no claim, $P_0 = \dfrac{e^{-\lambda}\lambda^0}{0!} = e^{-\lambda}$

The probability of one claim, $P_1 = \dfrac{e^{-\lambda}\lambda^1}{1!} = \lambda e^{-\lambda}$

The probability of two claims, $P_2 = \frac{e^{-\lambda}\lambda^2}{2!} = \frac{\lambda^2 e^{-\lambda}}{2}$

The probability of three claims, $P_3 = \frac{e^{-\lambda}\lambda^3}{3!} = \frac{\lambda^3 e^{-\lambda}}{6}$

The probability of four claims, $P_4 = \frac{e^{-\lambda}\lambda^4}{4!} = \frac{\lambda^4 e^{-\lambda}}{24}$

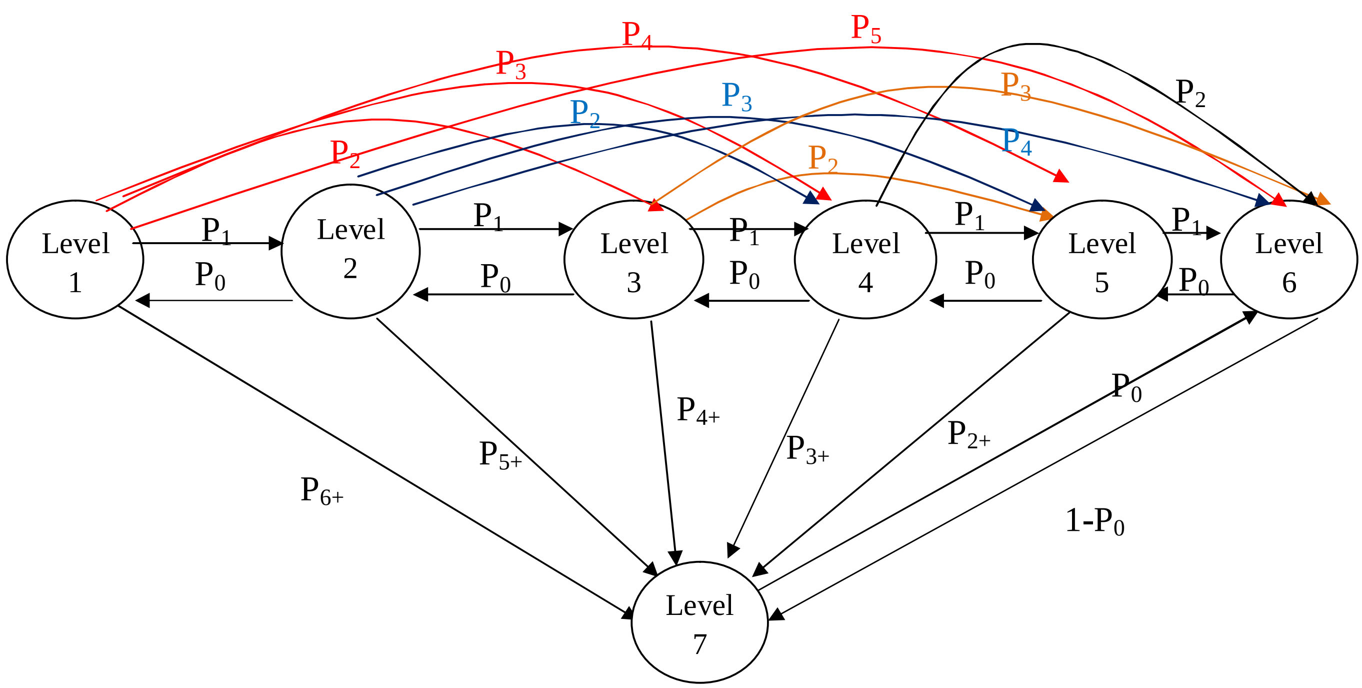
The probability of five claims, $P_5 = \frac{e^{-\lambda}\lambda^5}{5!} = \frac{\lambda^5 e^{-\lambda}}{120}$

The probability of six or more claims, P_6

$$= 1 - \left(e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2 e^{-\lambda}}{2} + \frac{\lambda^3 e^{-\lambda}}{6} + \frac{\lambda^4 e^{-\lambda}}{24} + \frac{\lambda^5 e^{-\lambda}}{120} \right)$$

$$= 1 - e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{6} + \frac{\lambda^4}{24} + \frac{\lambda^5}{120} \right)$$

Transition graph 3.6.1



Transition matrix 3.6.1

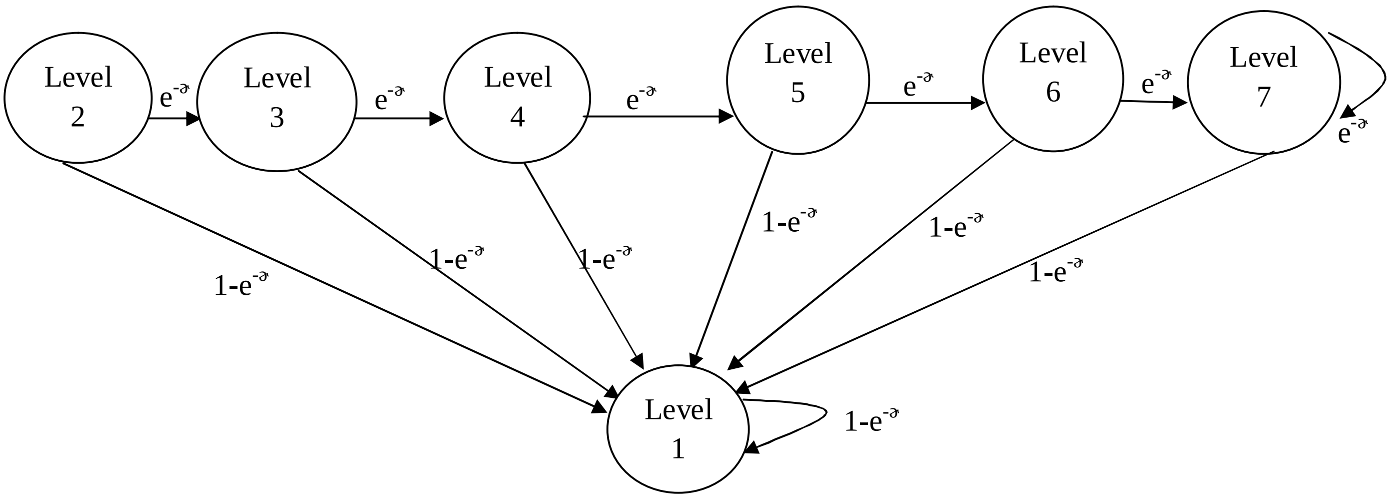
$$\begin{bmatrix} e^{-\lambda} & \lambda e^{-\lambda} & \frac{\lambda^2 e^{-\lambda}}{2} & \frac{\lambda^3 e^{-\lambda}}{6} & \frac{\lambda^4 e^{-\lambda}}{24} & \frac{\lambda^5 e^{-\lambda}}{120} & 1-e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{6}+\frac{\lambda^4}{24}+\frac{\lambda^5}{120}\right) \\ e^{-\lambda} & 0 & \lambda e^{-\lambda} & \frac{\lambda^2 e^{-\lambda}}{2} & \frac{\lambda^3 e^{-\lambda}}{6} & \frac{\lambda^4 e^{-\lambda}}{24} & 1-e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{6}+\frac{\lambda^4}{24}\right) \\ 0 & e^{-\lambda} & 0 & \lambda e^{-\lambda} & \frac{\lambda^2 e^{-\lambda}}{2} & \frac{\lambda^3 e^{-\lambda}}{6} & 1-e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{6}\right) \\ 0 & 0 & e^{-\lambda} & 0 & \lambda e^{-\lambda} & \frac{\lambda^2 e^{-\lambda}}{2} & 1-e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2}\right) \\ 0 & 0 & 0 & e^{-\lambda} & 0 & \lambda e^{-\lambda} & 1-e^{-\lambda}\left(1+\lambda\right) \\ 0 & 0 & 0 & 0 & e^{-\lambda} & 0 & 1-e^{-\lambda} \\ 0 & 0 & 0 & 0 & 0 & e^{-\lambda} & 1-e^{-\lambda} \end{bmatrix}$$

System 3

The Kenyan system has 7 classes as well with levels 40, 50, 60, 70, 80, 90 and 100 (starting level). When a policyholder makes a claim, he loses all discounts, while a claim free year leads to him going down a premium level / up a discount level.

Levels	Premium level	Discount level
1	100	0
2	90	10
3	80	20
4	70	30
5	60	40
6	50	50
7	40	60

Transition graph 3.6.2



The assumption is that, the number of claims conforms to a Poisson distribution, with parameter λ .

The probability of no claim $= e^{-\lambda}$

The probability of a claim $= 1 - e^{-\lambda}$

Transition matrix 3.6.2

$$\begin{bmatrix} 1-e^{-\lambda} & e^{-\lambda} & 0 & 0 & 0 & 0 & 0 \\ 1-e^{-\lambda} & 0 & e^{-\lambda} & 0 & 0 & 0 & 0 \\ 1-e^{-\lambda} & 0 & 0 & e^{-\lambda} & 0 & 0 & 0 \\ 1-e^{-\lambda} & 0 & 0 & 0 & e^{-\lambda} & 0 & 0 \\ 1-e^{-\lambda} & 0 & 0 & 0 & 0 & e^{-\lambda} & 0 \\ 1-e^{-\lambda} & 0 & 0 & 0 & 0 & 0 & e^{-\lambda} \\ 1-e^{-\lambda} & 0 & 0 & 0 & 0 & 0 & e^{-\lambda} \end{bmatrix}$$

System 4

Jean Lemaire (2013), a comparative analysis of most European and Japanese BMS, looks at the British and Swedish BMS. The British BMSôs conclusions class after claims is summarised below:

Class	Premium level	Class after claims			
		0	1	2	3 or more
7	100	6	7	7	7
6	75	5	7	7	7
5	65	4	6	7	7
4	55	3	5	7	7
3	45	2	5	7	7
2	40	1	4	6	7
1	35	1	4	6	7

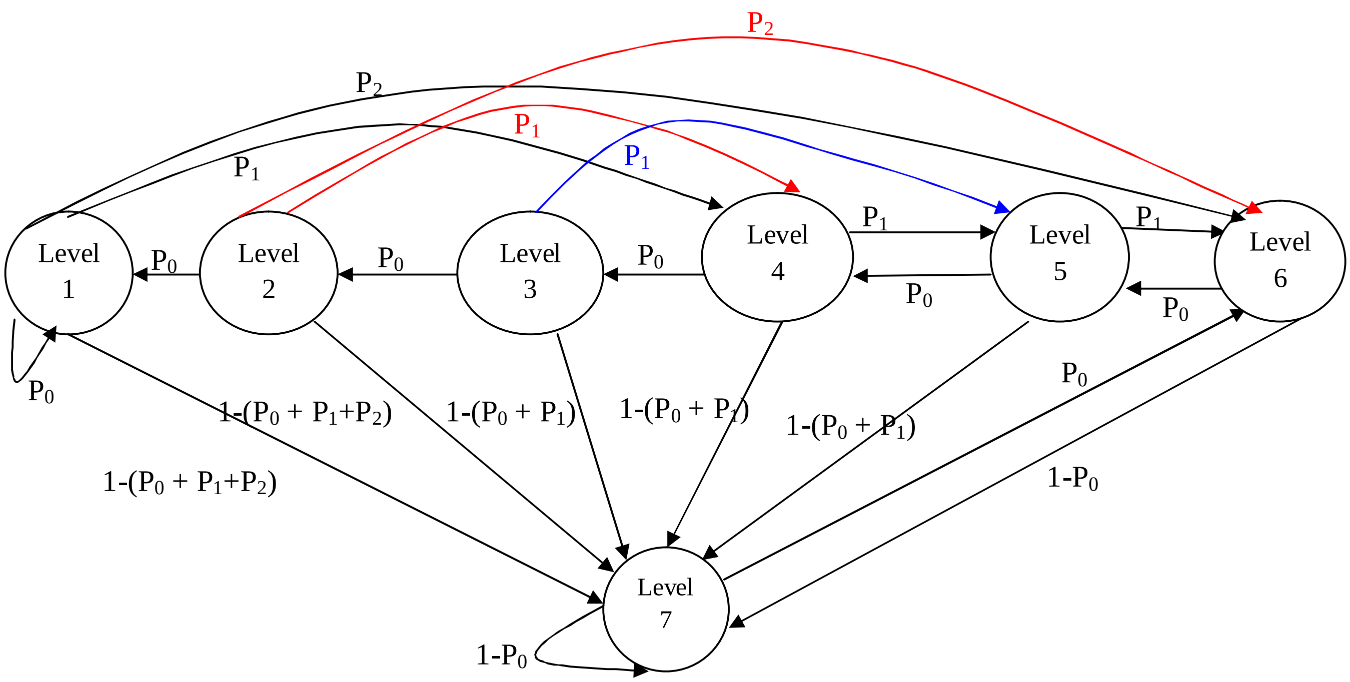
Let, the probability of no claim $= P_0$

The probability of a claim $= P_1$

The probability of two claims $= P_2$

The probability of three or more claims $= 1-(P_0 + P_1 + P_2)$

Transition graph 3.6.3



Transition matrix 3.6.3

$$\begin{bmatrix} P_0 & 0 & 0 & P_1 & 0 & P_2 & 1-(P_0+P_1+P_2) \\ P_0 & 0 & 0 & P_1 & 0 & P_2 & 1-(P_0+P_1+P_2) \\ 0 & P_0 & 0 & 0 & 0 & 0 & 1-(P_0+P_1) \\ 0 & 0 & P_0 & 0 & P_1 & 0 & 1-(P_0+P_1) \\ 0 & 0 & 0 & P_0 & 0 & P_1 & 1-(P_0+P_1) \\ 0 & 0 & 0 & 0 & P_0 & 0 & 1-P_0 \\ 0 & 0 & 0 & 0 & 0 & P_0 & 1-P_0 \end{bmatrix}$$

System 6

Jean Lemaire (2013) also looks at the Swedish BMS and the table below shows its transition rules.

Class	Premium level	Class after claims			
		$n=0$	$n=1$	$n=2$	$n\geq 3$
7	100	6	7	7	7
6	80	5	7	7	7
5	70	4	7	7	7
4	60	3	6	7	7
3	50	2	5	7	7
2	40	1 or 2	4	6	7
1	25	1	3	5	7

Let;

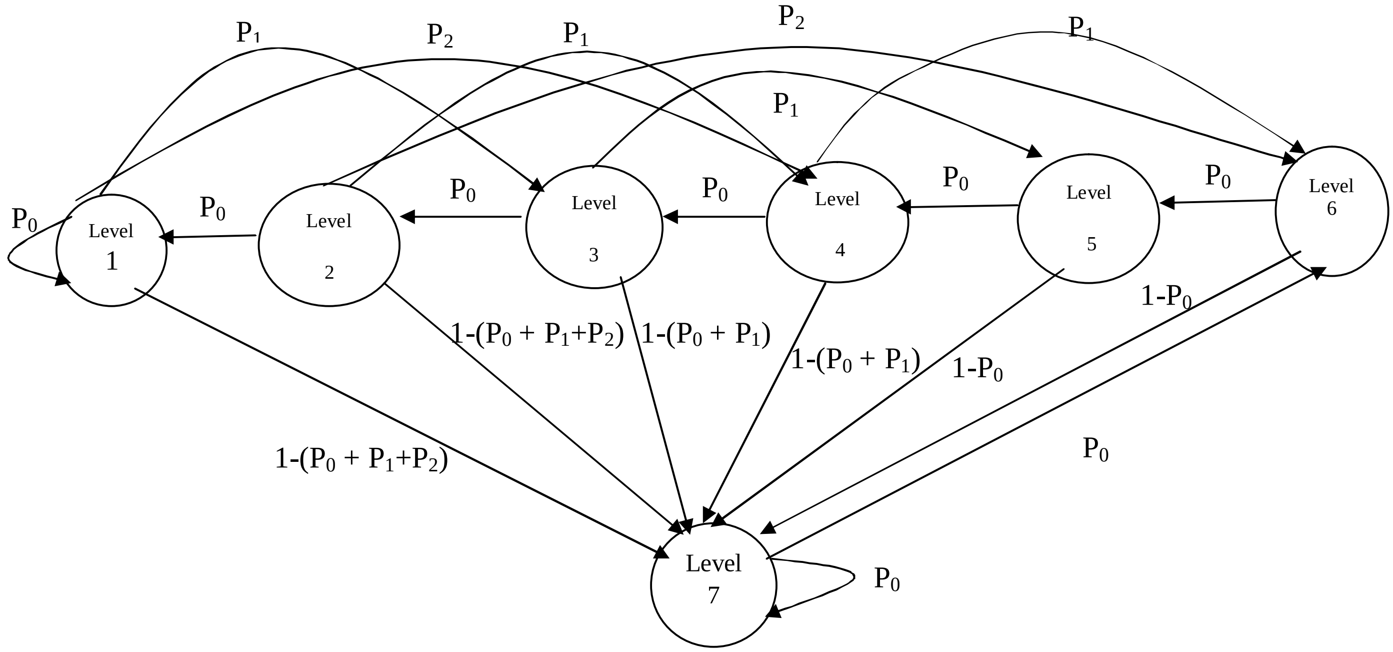
The probability of no claim $= P_0$

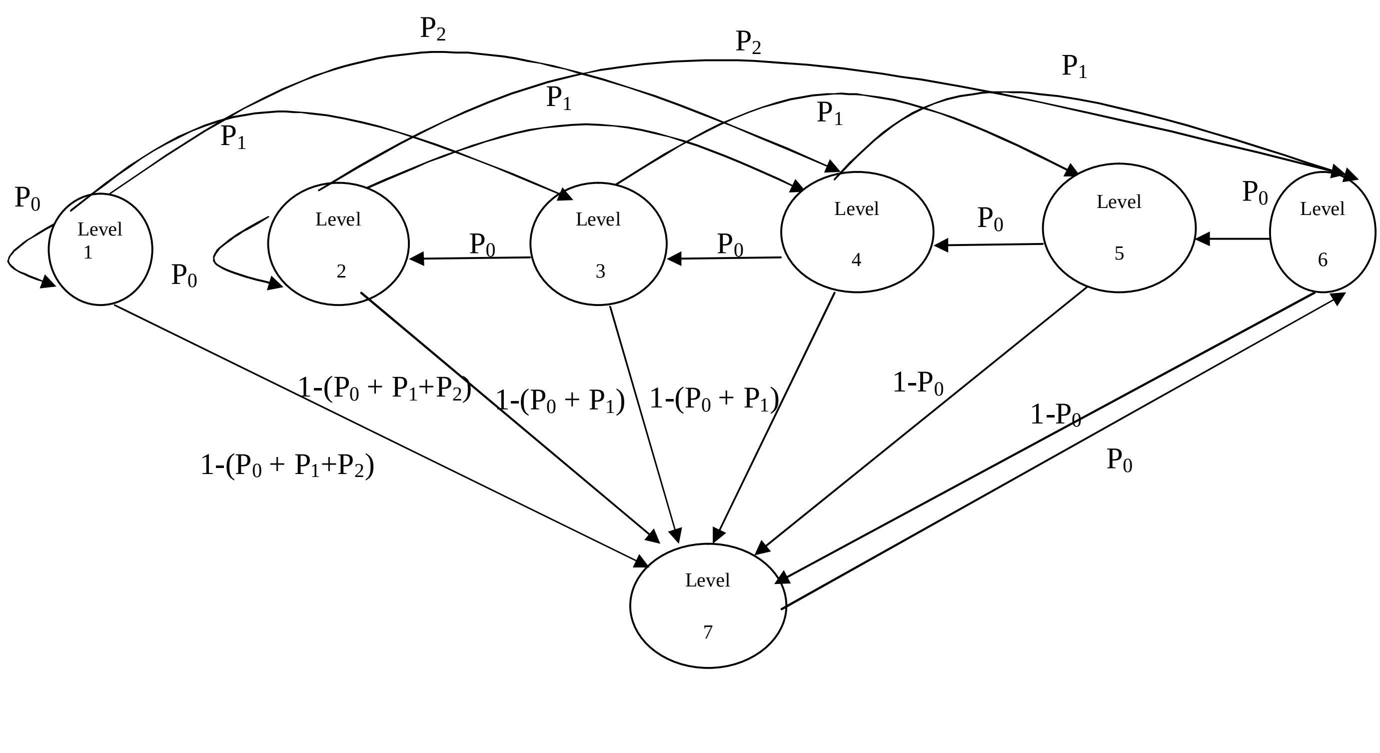
The probability of one claim $= P_1$

The probability of two claims $= P_2$

The probability of three or more claims $= 1-(P_0+P_1+P_2)$

<u>**Transition graph 3.6.4**</u>





Transition matrix 3.6.4

$$\begin{bmatrix} P_0 & 0 & P_1 & P_2 & 0 & 0 & 1-(P_0+P_1+P_2) \\ P_0 & 0 & 0 & P_1 & 0 & P_2 & 1-(P_0+P_1+P_2) \\ 0 & P_0 & 0 & 0 & P_1 & 0 & 1-(P_0+P_1) \\ 0 & 0 & P_0 & 0 & 0 & P_1 & 1-(P_0+P_1) \\ 0 & 0 & 0 & P_0 & 0 & 0 & 1-P_0 \\ 0 & 0 & 0 & 0 & P_0 & 0 & 1-P_0 \\ 0 & 0 & 0 & 0 & 0 & P_0 & 1-P_0 \end{bmatrix}$$

$$\begin{bmatrix} P_0 & 0 & P_1 & 0 & 0 & P_2 & 1-(P_0+P_1+P_2) \\ 0 & P_0 & 0 & P_1 & P_1 & 0 & 1-(P_0+P_1+P_2) \\ 0 & P_0 & 0 & 0 & 0 & P_1 & 1-(P_0+P_1) \\ 0 & 0 & P_0 & 0 & 0 & 0 & 1-(P_0+P_1) \\ 0 & 0 & 0 & P_0 & P_0 & 0 & 1-P_0 \\ 0 & 0 & 0 & 0 & 0 & P_0 & 1-P_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1-P_0 \end{bmatrix}$$

Remark

The Kenyan system is unfair and does not take into consideration the claim frequency.

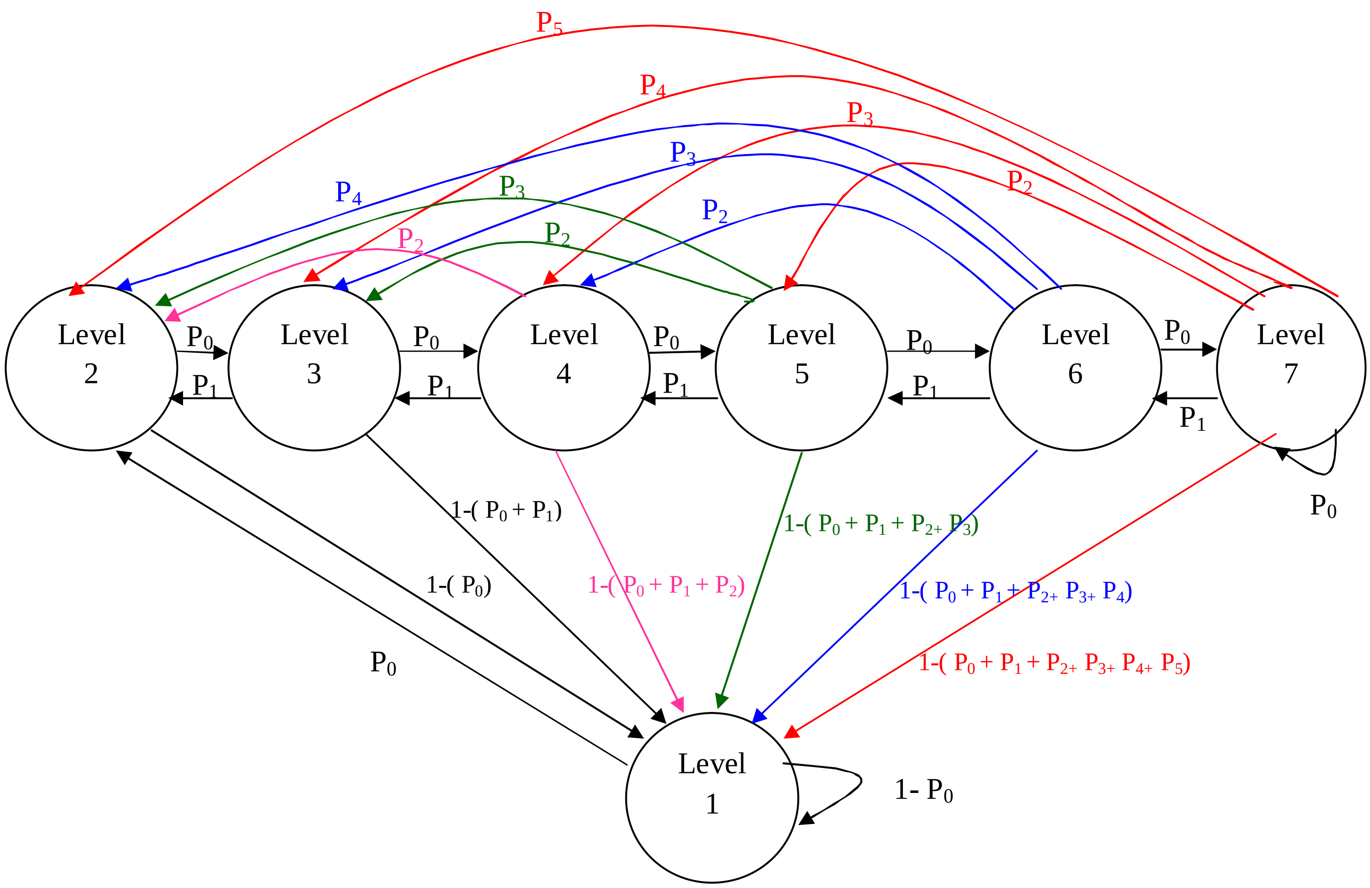
The Thailand BMS only gives favourable options to policyholders which is a disadvantage to the insurance companies.

An ideal seven stage BMS would be as follows:

Level	Level occupied if;						
	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n\geq 6$
7	7	6	5	4	3	2	1
6	7	5	4	3	2	1	1
5	6	4	3	2	1	1	1
4	5	3	2	1	1	1	1
3	4	2	1	1	1	1	1
2	3	1	1	1	1	1	1
1	2	1	1	1	1	1	1

Level 7 is the highest level of discount.

Transition graph 3.6.5



Transition matrix 3.6.5

$$
\begin{array}{c}
7\\ 6\\ 5\\ 4\\ 3\\ 2\\ 1
\end{array}
\begin{bmatrix}
P_0 & P_1 & P_2 & P_3 & P_4 & P_5 & 1-(P_0+P_1+P_2+P_3+P_4+P_5)\\
P_0 & 0 & P_1 & P_2 & P_3 & P_4 & 1-(P_0+P_1+P_2+P_3+P_4)\\
0 & P_0 & 0 & P_1 & P_2 & P_3 & 1-(P_0+P_1+P_2+P_3)\\
0 & 0 & P_0 & 0 & P_1 & P_2 & 1-(P_0+P_1+P_2)\\
0 & 0 & 0 & P_0 & 0 & P_1 & 1-(P_0+P_1)\\
0 & 0 & 0 & 0 & P_0 & 0 & 1-P_0\\
0 & 0 & 0 & 0 & 0 & P_0 & 1-P_0
\end{bmatrix}
$$

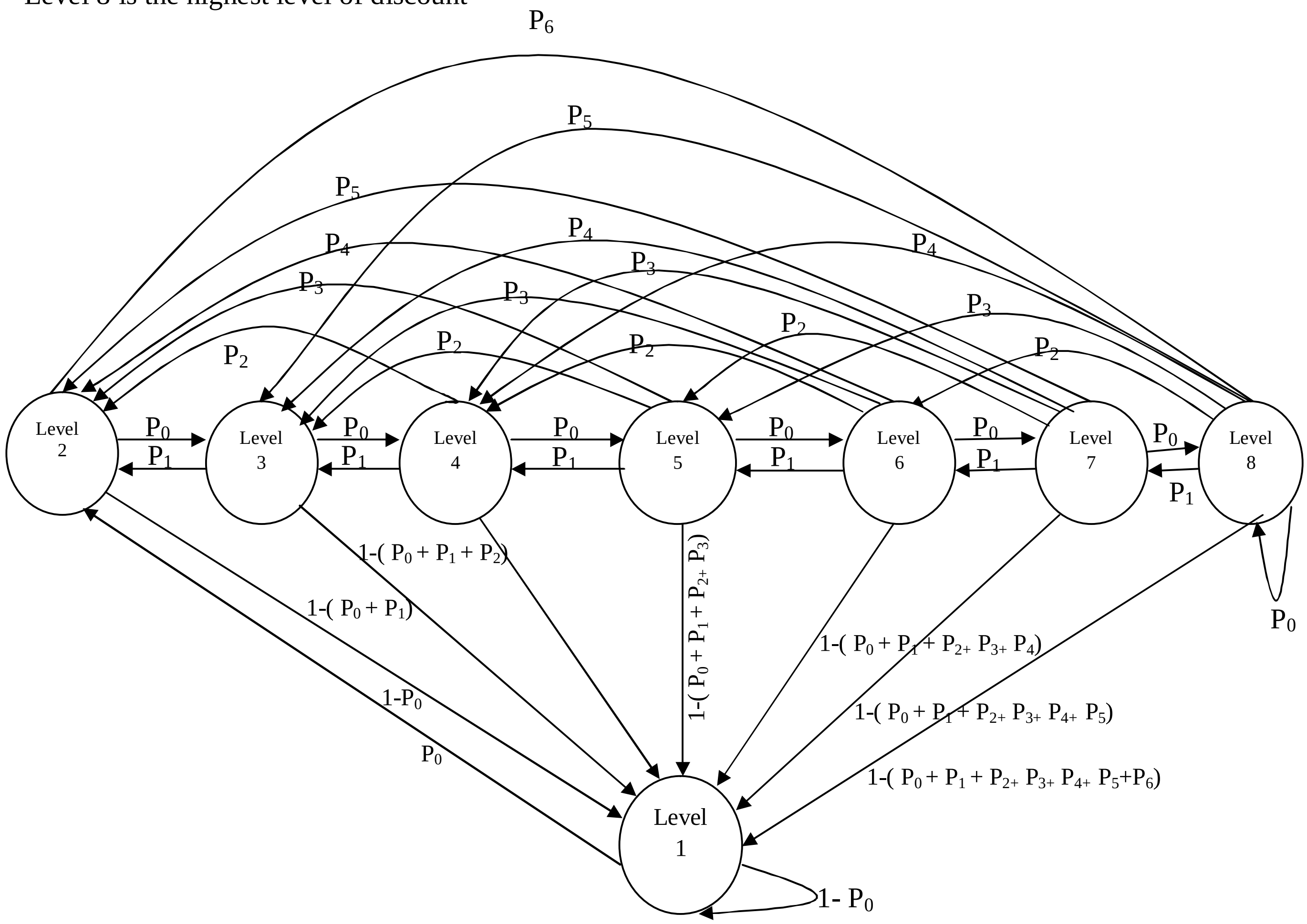
This is the same Brazilian BMS looked at by Lemaries (2013) in his article, North American Actuarial Journal

3.4. THE EIGHT STAGE NCD SYSTEM

An idea should have the following transition rule

Level	Level occupied if							
8	n=0	n=14	n=2	n=3	n=4	n=5	n=6	nÓ7
7	8	7	6	5	4	3	2	1
6	7	6	5	4	3	2	1	1
5	6	5	4	3	2	1	1	1
4	5	4	3	2	1	1	1	1
3	4	3	2	1	1	1	1	1
2	3	2	1	1	1	1	1	1
1	2	1	1	1	1	1	1	1

Level 8 is the highest level of discount



3.7 THE NINE STAGE BMS

J. F. Walhin and J. Paris (1999), give an illustration of a BMS with nine classes.

Class	0	1	2	3	4	5	6	7	8
Premium (percentage)	75	80	90	95	100	150	170	185	250

The transition rules are as follows:

-1 if no accident during the year.

$+3$ per accident during the year.

$+0$ if any accident in class 8.

-0 if no accident in class 0.

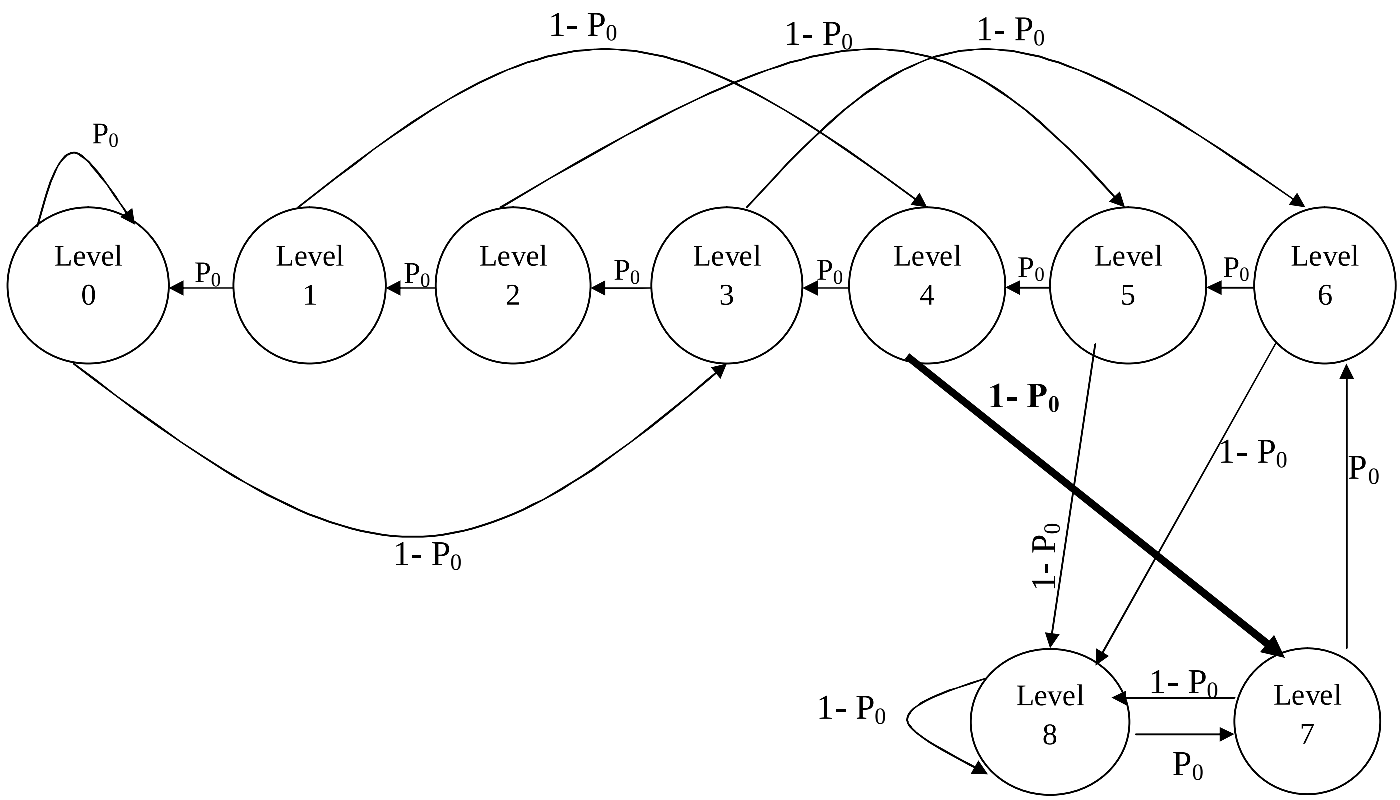
Class after claims can be summarised as follows:

Class	Premium level	Class after claims		Discount level
		$n=0$	$n\geq 1$	
8	250	7	8	-150
7	185	6	8	-85
6	170	5	8	-70
5	150	4	8	-50
4	100	3	7	0
3	95	2	6	5
2	90	1	5	10
1	80	0	4	20
0	75	0	3	25

Let;

The probability of no claim $=P_0$ and the probability of claim(s) $=1-P_0$

Transition graph 3.7.0



Transition matrix 3.7.0

$$\begin{array}{c} 8\\7\\6\\5\\4\\3\\2\\1\\0 \end{array}\begin{bmatrix} 1-P_0 & P_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1-P_0 & 0 & P_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1-P_0 & 0 & 0 & P_0 & 0 & 0 & 0 & 0 & 0 \\ 1-P_0 & 0 & 0 & 0 & P_0 & 0 & 0 & 0 & 0 \\ 0 & 1-P_0 & 0 & 0 & 0 & P_0 & 0 & 0 & 0 \\ 0 & 0 & 1-P_0 & 0 & 0 & 0 & P_0 & 0 & 0 \\ 0 & 0 & 0 & 1-P_0 & 0 & 0 & 0 & P_0 & 0 \\ 0 & 0 & 0 & 0 & 1-P_0 & 0 & 0 & 0 & P_0 \\ 0 & 0 & 0 & 0 & 0 & 1-P_0 & 0 & 0 & P_0 \end{bmatrix}$$

Remark

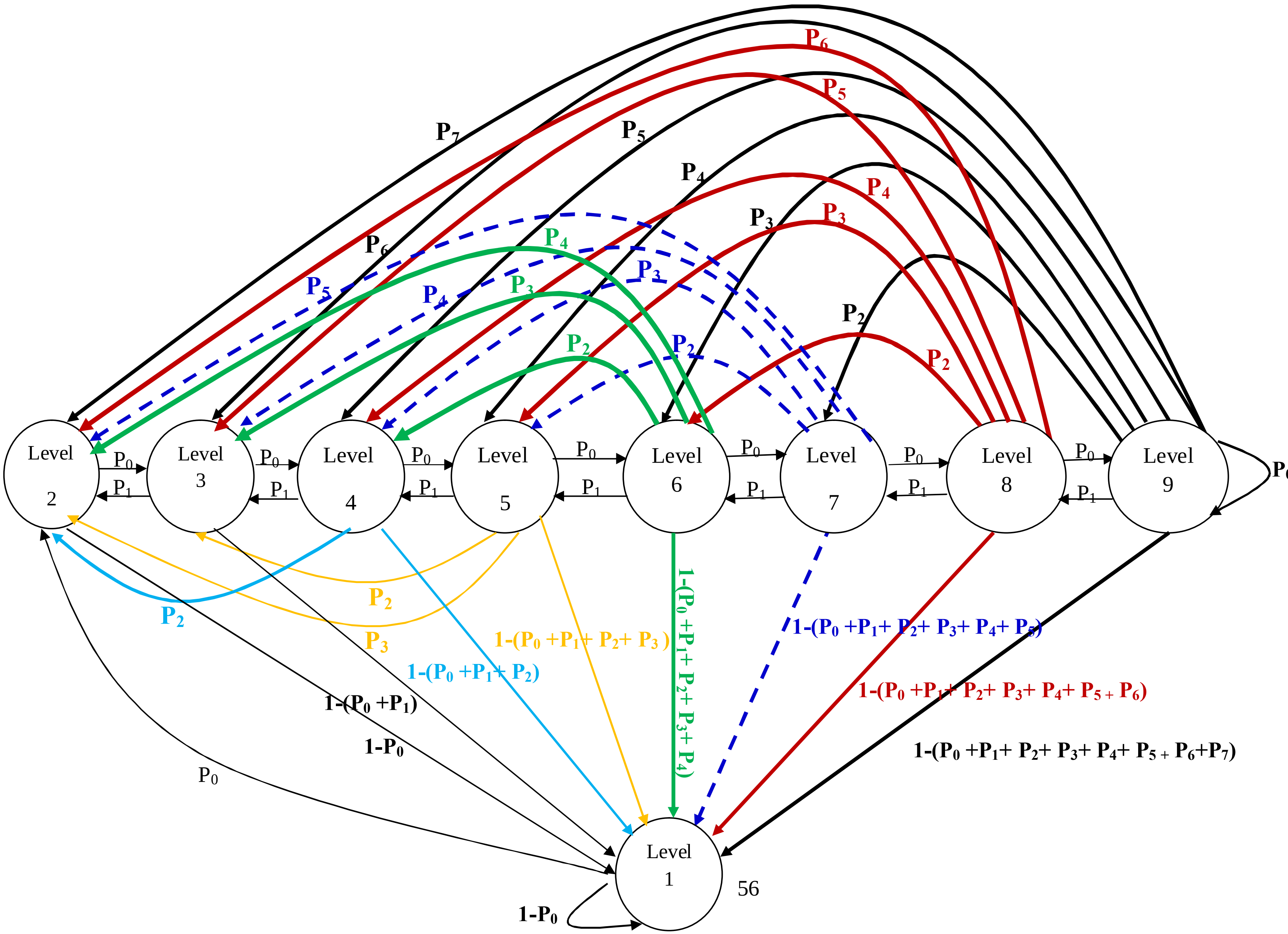
The nine stage BMS illustrated by J. F. Walhin and J. Paris (1999) does not consider the number of claims reported by individuals.

The ideal nine stage BMS that considers claims frequency is shown by the table below;

Level	Level occupied if								
	$n = o$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n \geq 8$
9	9	8	7	6	5	4	3	2	1
8	9	7	6	5	4	3	2	1	1
7	8	6	5	4	3	2	1	1	1
6	7	5	4	3	2	1	1	1	1
5	6	4	3	2	1	1	1	1	1
4	5	3	2	1	1	1	1	1	1
3	4	2	1	1	1	1	1	1	1
2	3	1	1	1	1	1	1	1	1
1	2	1	1	1	1	1	1	1	1

Level 9 is the highest level of discount.

Transition graph 3.7.1



Transition matrix 3.7.1

$$\begin{matrix} 9 \\ 8 \\ 7 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{matrix}\begin{bmatrix} P_0 & P_1 & P_2 & P_3 & P_4 & P_5 & P_6 & P_7 & 1-(P_0+P_1+P_2+P_3+P_4+P_5+P_6+P_7) \\ P_0 & 0 & P_1 & P_2 & P_3 & P_4 & P_5 & P_6 & 1-(P_0+P_1+P_2+P_3+P_4+P_5+P_6) \\ 0 & P_0 & 0 & P_1 & P_2 & P_3 & P_4 & P_5 & 1-(P_0+P_1+P_2+P_3+P_4+P_5) \\ 0 & 0 & P_0 & 0 & P_1 & P_2 & P_3 & P_4 & 1-(P_0+P_1+P_2+P_3+P_4) \\ 0 & 0 & 0 & P_0 & 0 & P_1 & P_2 & P_3 & 1-(P_0+P_1+P_2+P_3) \\ 0 & 0 & 0 & 0 & P_0 & 0 & P_1 & P_2 & 1-(P_0+P_1+P_2) \\ 0 & 0 & 0 & 0 & 0 & P_0 & 0 & P_1 & 1-(P_0+P_1) \\ 0 & 0 & 0 & 0 & 0 & 0 & P_0 & 0 & 1-P_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & P_0 & 1-P_0 \end{bmatrix}$$

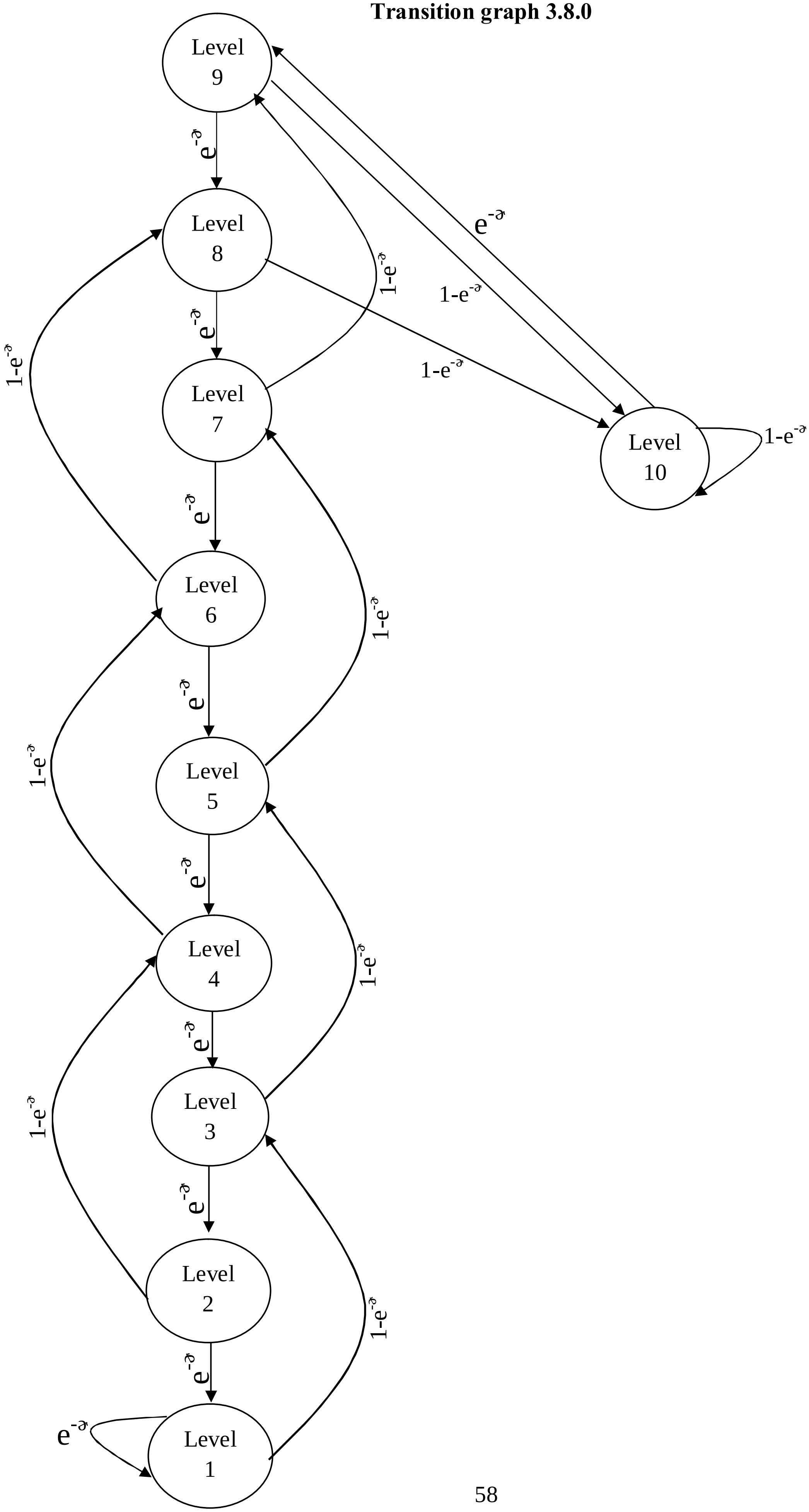
3.9. THE TEN STAGE BMS

Lemaire and Zi (1994) look at the automobile third party insurance merit-rating systems of 22 countries among them is Denmark, whose BMS has 10 number of classes with levels 30, 40, 50, 60, 70, 80, 90, 100, 120, and 150 with level 100 as the starting level. For a claim free year, a policyholder moves down a premium level and for a year with claim(s), he moves up two levels.

Class	Premium level	Discount level	Class after claims	
			$n=0$	$n\geq 1$
10	150	-50	9	10
9	120	-20	8	10
8	100	0	7	10
7	90	10	6	9
6	80	20	5	8
5	70	30	4	7
4	60	40	3	6
3	50	50	2	5
2	40	60	1	4
1	30	70	1	3

The probability of no claim $= e^{-\lambda}$ and the probability of claim(s) $=1-e^{-\lambda}$

Transition graph 3.8.0



Transition matrix 3.8.0

$$
\begin{array}{c}
1\\2\\3\\4\\5\\6\\7\\8\\9\\10
\end{array}
\begin{bmatrix}
e^{-\lambda} & 0 & 1-e^{-\lambda} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & e^{-\lambda} & 0 & 0 & 1-e^{-\lambda} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & e^{-\lambda} & 0 & 1-e^{-\lambda} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e^{-\lambda} & 1-e^{-\lambda}
\end{bmatrix}
$$

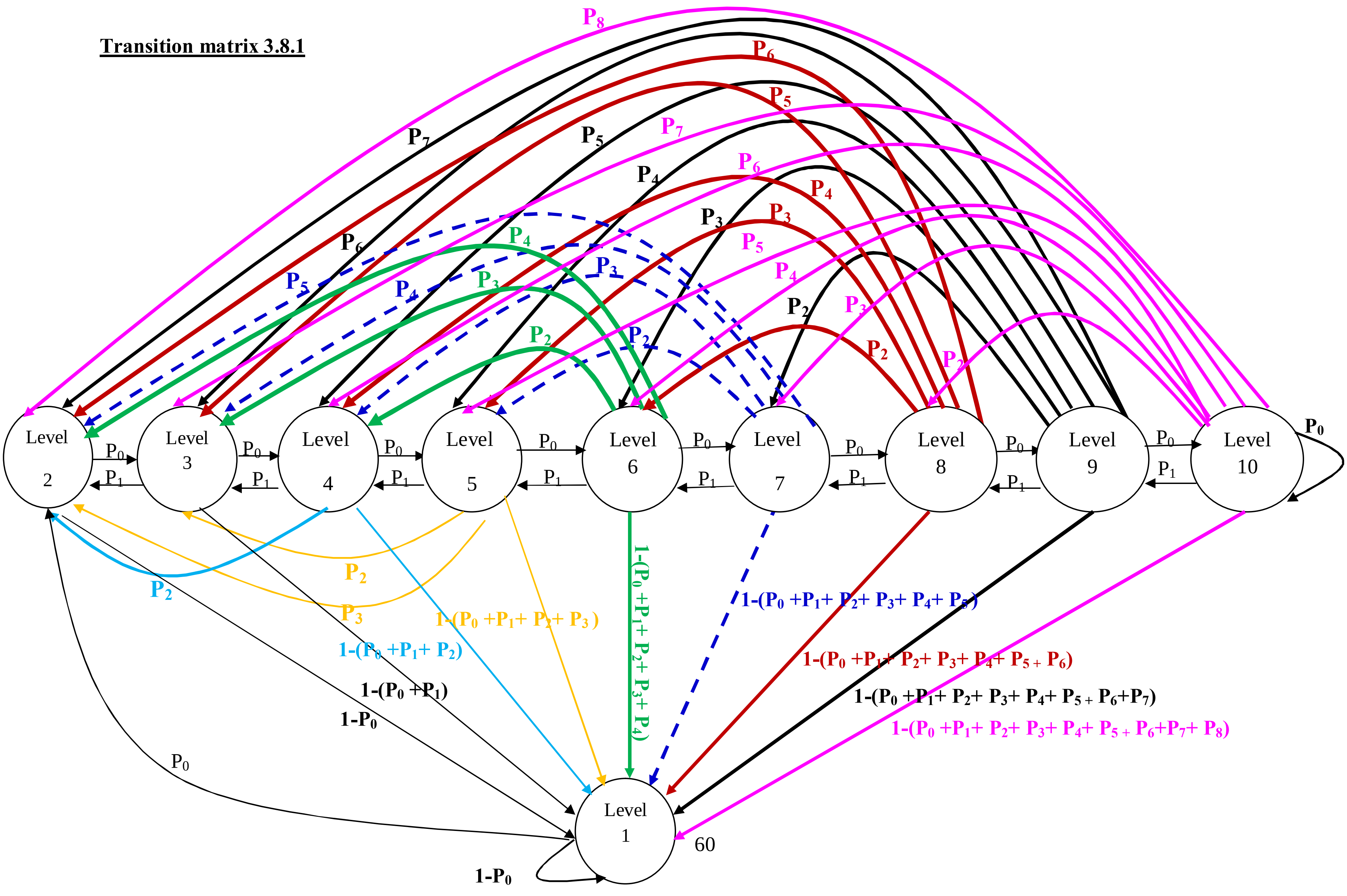
Remark

The table below represents the transition rules for an ideal ten stage BMS.

Level	Level occupied if									
	$n=o$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n\geq 9$
10	10	9	8	7	6	5	4	3	2	1
9	10	8	7	6	5	4	3	2	1	1
8	9	7	6	5	4	3	2	1	1	1
7	8	6	5	4	3	2	1	1	1	1
6	7	5	4	3	2	1	1	1	1	1
5	6	4	3	2	1	1	1	1	1	1
4	5	3	2	1	1	1	1	1	1	1
3	4	2	1	1	1	1	1	1	1	1
2	3	1	1	1	1	1	1	1	1	1
1	2	1	1	1	1	1	1	1	1	1

Level 10 is the highest discount level

Transition matrix 3.8.1



Transition matrix 3.8.1

$$\begin{array}{c} 10 \\ 9 \\ 8 \\ 7 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{array}\begin{bmatrix} P_0 & P_1 & P_2 & P_3 & P_4 & P_5 & P_6 & P_7 & P_8 & 1-(P_0+P_1+P_2+P_3+P_4+P_5+P_6+P_7+P_8) \\ P_0 & 0 & P_1 & P_2 & P_3 & P_4 & P_5 & P_6 & P_7 & 1-(P_0+P_1+P_2+P_3+P_4+P_5+P_6+P_7) \\ 0 & P_0 & 0 & P_1 & P_2 & P_3 & P_4 & P_5 & P_6 & 1-(P_0+P_1+P_2+P_3+P_4+P_5+P_6) \\ 0 & 0 & P_0 & 0 & P_1 & P_2 & P_3 & P_4 & P_5 & 1-(P_0+P_1+P_2+P_3+P_4+P_5) \\ 0 & 0 & 0 & P_0 & 0 & P_1 & P_2 & P_3 & P_4 & 1-(P_0+P_1+P_2+P_3+P_4) \\ 0 & 0 & 0 & 0 & P_0 & 0 & P_1 & P_2 & P_3 & 1-(P_0+P_1+P_2+P_3) \\ 0 & 0 & 0 & 0 & 0 & P_0 & 0 & P_1 & P_2 & 1-(P_0+P_1+P_2) \\ 0 & 0 & 0 & 0 & 0 & 0 & P_0 & 0 & P_1 & 1-(P_0+P_1) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & P_0 & 0 & 1-P_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P_0 & 1-P_0 \end{bmatrix}$$

CHAPTER IV

CLAIM FREQUENCY DISRIBUTION

4.1 INTRODUCTION

Poisson distribution is commonly used in describing the probability of the number of claims within a given period (claim frequency distribution). The distribution has a constant parameter, that is, the characteristics of policyholders (drivers) are assumed to be homogeneous. This is not the case in reality. To take care of the non-homogeneity factor, most studies are now using the mixed Poisson distributions to describe the claim frequency distribution.

Poisson mixtures capture much of the heterogeneous structure and they fit data better than the standard Poisson. In this chapter, we are going to consider mixed Poisson distributions using the following mixing distributions:

i. Exponential
ii. One parameter gamma.
iii. Two parameter gamma.
iv. Lindley.

For each mixed distribution, we are going to estimate its parameters. The estimation methods we will look at include:

- The method of moments.
- The maximum likelihood method.

4.2 PARAMETER ESTIMATION

4.2.1 THE METHOD OF MOMENTS

To obtain the method of moment estimator (MME), first we derive the equations that relate the population moments to the parameters of interest, that is, the expected values. A sample is then drawn and the population moments are estimated from the sample. The equations are then solved for the parameters of interest, using the sample moments in place of the unknown population moments. This will result in the estimates of those parameters.

Suppose the problem is to estimate, k, unknown parameters, $\theta_1, \cdots, \theta_k$ with the distribution characterized as $f_W(w;\theta)$.

Suppose the first k moments of the true distribution (the population moments) can be expressed as functions of θs :

$$\mu_1 = E[W] = g_1(\theta_1, \cdots, \theta_k)$$
$$\mu_2 = E[W^2] = g_2(\theta_1, \cdots, \theta_k)$$
$$\vdots$$
$$\mu_k = E[W^k] = g_k(\theta_1, \cdots, \theta_k)$$

Suppose a sample of size n is drawn, resulting in the values $w_1, \cdots, w_n$ for $j = 1, \cdots, k$

Let $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} w_j^{\,i}$ be the $j-th$ sample moment, an estimate of μ_j .

The method of moments estimator for $\hat{\theta}_1, \cdots, \hat{\theta}_k$ is defined as the solution to the equations:

$$\hat{\mu}_1 = g_1(\hat{\theta}_1, \cdots, \hat{\theta}_k)$$
$$\hat{\mu}_2 = g_2(\hat{\theta}_1, \cdots, \hat{\theta}_k)$$
$$\vdots$$
$$\hat{\mu}_k = g_k(\hat{\theta}_1, \cdots, \hat{\theta}_k)$$

4.2.2 MAXIMUM LIKELIHOOD METHOD

It is a method of estimating the parameters of a statistical model. For a fixed set of data and underlying statistical model, this method selects the set of values of the model parameter that maximizes the likelihood function.

Let $f(x_1, \cdots, x_n; \theta)$ be the joint probability or density function of n , random variable $X_1, \cdots, X_n$ with sample values $x_1, \cdots, x_n$. The likelihood function of the sample is given by;

$$L(\theta; x_1, \cdots, x_n) = f(x_1, \cdots, x_n)$$

$L(\theta)$ is a briefer notation

L is a function of θ for a fixed sample values

If $X_1, \cdots, X_n$ are discrete iid random variables with probability function $p(x, \theta)$ then the likelihood function is given by;

$$L(\theta) = P(X_1 = x_1, \cdots, X_n = x_n)$$

In the continuous case, if density is $f(x, \theta)$, then the likelihood function is;

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

$$= \prod_{i=1}^{n} P(X_i = x_i)$$

$$= \prod_{i=1}^{n} P(x_i;\theta)$$

Although the likelihood function depends on the observed sample value, $x = (x_1, \cdots, x_n)$, it is regarded as a function of parameter θ.

4.3 GAMMA FUNCTIONS

Definition 1:

$$(i)\Gamma_{(\alpha)} = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

$$(ii)\Gamma_{(\alpha+1)} = \int_0^\infty t^{\alpha} e^{-t} dt$$

Property 1.1

Integration by parts of definition 1(ii);

$$\int f(x)g(x)dx = f(x)G(x) - \int f'(x)G(x)dx$$

Let $f(x) = t^\alpha \Rightarrow f'(x) = \alpha t^{\alpha-1}$ and $g(x) = e^{-t} \Rightarrow -e^{-t}$

Thus;

$$\Gamma(\alpha+1) = \int_0^\infty t^\alpha e^{-t} dt$$

$$= -t^\alpha e^{-t} \Big|_0^\infty - \int_0^\infty \alpha t^{\alpha-1} \bullet -e^{-t} dt$$

$$= \alpha \int_0^\infty t^{\alpha-1} e^{-t} dt$$

$$= \alpha\Gamma(\alpha)$$

$$\boxed{\Gamma(\alpha+1) = \alpha\Gamma(\alpha)}$$

Property 1.2

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t}\, dt$$

$$\Gamma(1) = \int_0^{\infty} t^{1-1} e^{-t}\, dt$$

$$= \int_0^{\infty} e^{-t}\, dt$$

$$= -e^{-t}\Big|_0^{\infty} = -(0-1) = 1$$

$$\boxed{\Gamma(1) = 1}$$

Property 1.3

$$\Gamma(n+1) = n\Gamma(n)$$

$$= n(n-1)\Gamma(n-1)$$

$$= n(n-1)(n-2)\Gamma(n-2)\cdots$$

$$= n(n-1)(n-2)\cdots 2 \bullet 1 \bullet \Gamma(1)$$

$$= n!$$

$\Gamma\left(n+1\right) = n!$, n is a positive integer.

Property 1.4

$$\int_0^{\infty} e^{-\beta x} x^{\alpha}\, dx = \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} \ ; \ \alpha, \beta \geq 0$$

Proof;

Let $y = \beta x \Rightarrow x = \frac{y}{\beta}$ meaning $\frac{dy}{dx} = dx$

Therefore;

$$\int_0^{\infty} e^{-\beta x} x^{\alpha}\, dx = \int_0^{\infty} e^{-y} \left(\frac{y}{\beta}\right)^{\alpha} {dy}/{\beta}$$

$$= \frac{1}{\beta^{\alpha} \times \beta} \int_0^{\infty} e^{-y} y^{\alpha} dy$$

$$= \frac{1}{\beta^{\alpha+1}} \Gamma(\alpha+1)$$

$$= \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}}$$

4.4 MIXED POISSON DISTRIBUTIONS

4.4.1 INTRODUCTION

In this chapter, we are going to use mixed Poisson distributions when the mixing distributions are, exponential, one parameter and two parameter gamma and the Lindley distributions.

The parameters of these distributions have been estimated using the method of moments and the maximum likelihood method.

The following section will briefly describe the estimation of parameters.

We have also considered each mixed distribution separately, finding its mean and variance and estimating their parameters.

A probability distribution is said to be a mixed distribution if its probability density function can be expressed as;

$$f(x) = \int_{\Theta} f(x/\lambda) g_{\lambda}(\lambda) d\lambda \ , \ \lambda \in \Theta$$

Where;

$f(x/\lambda)g(\lambda) =$ mixture or mixed distribution

$f(x/\lambda) =$ conditional distribution

$g(\lambda) =$ mixing density / distribution

Regardless of the form of $f(x/\lambda)$, the expected value of a function $h(x)$ is obtained as;

$$E[h(x)] = \int_{\Theta} E_{x/\lambda}[h(x)] g(\lambda) d\lambda$$

$$E[X] = E[E_{x/\Lambda}] = E[\Lambda]$$

The variance of X in the mixed distribution is the sum of the variance of its conditional mean and the mean of its conditional variance.

$$V[X] = V[E_{x/\lambda}] + E[V_{x/\lambda}] \cdots (i)$$

A random variable X follows a mixed Poisson distribution with mixing distribution having a probability density function, g , if its probability function is;

$$P(X=x)=p(x)=\int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}g(\lambda)d\lambda \quad ,x=0,1,\cdots$$

In terms of probability generating function, $H(S)$ of X:

$$H(S)=\int_0^\infty e^{-\lambda(S-1)}g(\lambda)d\lambda\cdots(ii)$$

$M_X(S-1)$ is the moment generating function of a mixed distribution evaluated at $S-1$. Meaning the probability generating function of a mixed Poisson distribution determines the mixing distribution through its moment generating function.

Consider a random variable, X, whose distribution is a mixed Poisson distribution. Then;

$$(i)P(X\le x)=\int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}G(\lambda)d\lambda$$

$$(ii)P(X=x)=\int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}[1-G(\lambda)]d\lambda$$

Where $G(\lambda)=\int_0^\lambda g(x)dx$ is the distribution function of parameter λ .

The variance of the mixed Poisson distribution:

$$\begin{aligned}Var(X)&=E[X^2]-(E[X])^2\\&=E[\lambda^2]+E[\lambda]-(E[X])^2\\&=E[\lambda]+Var(\lambda)\end{aligned}$$

For a probability generating function $Q(t)$:

$$Q(t)=E[t^X]=\int_0^\infty \exp[\lambda(t-1)]g(\lambda)d\lambda$$

Factorial moments of the mixed Poisson distribution are the same as the moments of the mixing distribution about the origin. Thus, the moments about the origin of the mixed Poisson distribution can be expressed in terms of those of the mixing distribution.

4.4.2 POISSON – EXPONENTIAL MIXTURE (GEOMETRIC DISTRIBUTION)

Assuming that the number of claims, x, is Poisson distributed with parameter $\lambda > 0$, and λ is distributed with parameter μ.

$$f(x/\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}; x = 0,1,\cdots \text{ and } \lambda > 0$$

The distribution of λ is;

$$g(\lambda;\mu) = \mu e^{-\mu\lambda} d\lambda \quad ;\mu,\lambda > 0$$

The mixture:

$$f(x) = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!} \bullet \mu e^{-\mu\lambda} d\lambda$$

$$= \frac{\mu}{x!}\int_0^\infty e^{-\lambda}\lambda^x \bullet e^{-\mu\lambda} d\lambda$$

$$= \frac{\mu}{x!}\int_0^\infty e^{-\lambda(1+\mu)}\lambda^x d\lambda$$

$$= \frac{\mu}{x!} \bullet \frac{\Gamma(x+1)}{(1+\mu)^{x+1}}$$

$$= \frac{\mu}{x!} \bullet \frac{x!}{(1+\mu)^{x+1}} = \frac{\mu}{(1+\mu)^{x+1}}$$

$$f(x) = \frac{\mu}{1+\mu} \bullet \frac{1}{(1+\mu)^x}$$

which is a Geometric distribution.

In terms of pgf ;

$$H(S) = \int_0^\infty e^{-\lambda(S-1)} \mu e^{-\mu\lambda} d\lambda$$

$$= \mu\int_0^\infty e^{-\lambda[(-S+1)+\mu]} d\lambda$$

$$= \mu\left[\frac{e^{-\lambda[(-S+1)+\mu]}}{(1-S)+\mu}\right]_0^\infty$$

$$= \mu\left[\frac{1}{(1-S)+\mu}\right] = \frac{\mu}{1+\mu-S}$$

$$= \frac{\frac{\mu}{1+\mu}}{1-\left(\frac{S}{1+\mu}\right)}$$

which is the pgf of a Geometric distribution with parameter $\frac{\mu}{1+\mu}$

The mean of the Geometric distribution – $H'(1)$

$$H(S) = \frac{\frac{\mu}{1+\mu}}{1-\left(\frac{S}{1+\mu}\right)} \Rightarrow \frac{\mu}{1+\mu} \times \left[1-\frac{S}{1+\mu}\right]^{-1}$$

$$H'(S) = \frac{\frac{\mu}{(1+\mu)^2}}{\left[1-\left(\frac{S}{1+\mu}\right)\right]^2}$$

$$H'(1) = \frac{\frac{\mu}{(1+\mu)^2}}{\left[1-\left(\frac{1}{1+\mu}\right)\right]^2} = \frac{\frac{\mu}{(1+\mu)^2}}{\left(\frac{\mu}{1+\mu}\right)^2} = \frac{1}{\mu}$$

$$\bar{X} = \frac{1}{\mu}$$

The variance of the Geometric distribution

$$H''(S) = \frac{\partial}{\partial S} H'(S)$$

$$H''(S) = \frac{\partial}{\partial S} \frac{\frac{\mu}{(1+\mu)^2}}{\left[1-\left(\frac{S}{1+\mu}\right)\right]^2} = \frac{2\mu}{(1+\mu)^3}\left[1-\left(\frac{S}{1+\mu}\right)\right]^{-3}$$

$$H''(1) = \frac{2\mu}{(1+\mu)^3} \bullet \left(\frac{\mu}{1+\mu}\right)^{-3} = \frac{2}{\mu^2}$$

$$Var[x] = H''(1) + H'(1) - [H'(1)]^2$$

$$Var[x] = \frac{2}{\mu^2} + \frac{1}{\mu} - \frac{1}{\mu^2} = \frac{1+\mu}{\mu^2}$$

Estimation of parameter, μ, of the Geometric distribution

i. **The method of moments**

$$E[x] = \bar{X} = \frac{1}{\mu}$$

$$\mu = \frac{1}{\bar{X}}$$

Thus the MME of μ is $\breve{\mu} = \frac{1}{\bar{X}}$

ii. **The maximum likelihood method**

$$L(\mu) = \prod_{i=1}^{n} \left(\frac{\mu}{1+\mu}\right)\left(\frac{1}{1+\mu}\right)^{x}$$

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

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

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

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

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

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

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

$$\sum_{i=1}^{n} x_i = \frac{n}{\mu}$$

$$\hat{\mu} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{X}}$$

4.4.3 POISSON-GAMMA MIXTURE (NEGATIVE BINOMIAL DISTRIBUTION)

4.4.3.1 The one parameter gamma mixing distribution

Let $f(x/\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}$; $x = 0,1,2,\cdots$

And the distribution of λ, $g(\lambda) = \frac{e^{-\lambda}\lambda^{\alpha-1}}{\Gamma(\alpha)}$

The mixture;

$$f(x) = \int_0^\infty \frac{e^{-\lambda}\lambda^{x}}{x!} \bullet \frac{e^{-\lambda}\lambda^{\alpha-1}}{\Gamma(\alpha)} d\lambda$$

$$= \frac{1}{x!\Gamma(\alpha)} \int_0^\infty e^{-2\lambda} \bullet \lambda^{x+\alpha-1} d\lambda$$

$$= \frac{1}{x!\Gamma(\alpha)} \bullet \frac{\Gamma(x+\alpha)}{2^{x+\alpha}} = \frac{\Gamma(x+\alpha)}{x!\Gamma(\alpha)} \bullet \frac{1}{2^{x+\alpha}}$$

$$f(x) = \binom{x+\alpha-1}{x} \left(\frac{1}{2}\right)^{x} \left(\frac{1}{2}\right)^{\alpha} ; x = 0,1,2,\cdots$$

which is a NBD with parameter Ŭ

In terms of pdf,

$$H(S) = \int_0^\infty e^{-\lambda(S-1)} g(\lambda) d\lambda$$

$$= \int_0^\infty e^{-\lambda(1-S)} \frac{e^{-\lambda}\lambda^{\alpha-1}}{\Gamma(\alpha)} d\lambda$$

$$= \frac{1}{\Gamma(\alpha)} \int_0^\infty e^{-\lambda(-S+2)} \lambda^{\alpha-1} d\lambda$$

$$= \frac{1}{\Gamma(\alpha)} \bullet \frac{\Gamma(\alpha)}{(2-S)^{\alpha}}$$

$$= \frac{1^{\alpha}}{(2-S)^{\alpha}} = \frac{0.5^{\alpha}}{(1-0.5S)^{\alpha}}$$

which is the pgf of NBD with parameter Ŭ

The mean of the one parameter gamma mixing distribution

$$E[X] = H'(1)$$

$$H'(S) = \frac{\partial}{\partial S} H(S)$$

$$= \frac{\partial}{\partial S} \frac{0.5^{\alpha}}{(1-0.5S)^{\alpha}}$$

$$= \frac{0.5\alpha \times 0.5^{\alpha}}{(1-0.5S)^{\alpha+1}}$$

$$H'(1) = \alpha = \bar{X}$$

The variance of the one parameter gamma mixing distribution

$$H''(S) = \frac{\partial}{\partial S} H'(S)$$

$$= \frac{\partial}{\partial S} \frac{0.5\alpha \times 0.5^{\alpha}}{(1-0.5S)^{\alpha+1}}$$

$$= \frac{0.5\alpha \bullet 0.5^{\alpha}[0.5\alpha + 0.5)}{(1-0.5S)^{\alpha+2}}$$

$$H''(1) = \frac{\alpha[0.5\alpha + 0.5]}{0.5}$$

$$Var[X] = H''(1) + H'(1) - [H'(1)]^2$$

$$= \frac{\alpha[0.5\alpha + 0.5]}{0.5} + \alpha - \alpha^2$$

$$= \frac{0.5\alpha^2 + 0.5\alpha + 0.5\alpha - 0.5\alpha^2}{0.5}$$

$$= \frac{\alpha}{0.5}$$

$$= 2\alpha$$

Estimation of the parameter α, of the one parameter gamma mixing distribution

a. **The method of moments**

$$\bar{X} = \alpha$$

$$\hat{\alpha} = \bar{X}$$

b. The maximum likelihood method

$$L(\alpha) = \prod_{i=1}^{n} \binom{x+\alpha-1}{x} \left(\frac{1}{2}\right)^{x} \left(\frac{1}{2}\right)^{\alpha}$$

$$= \prod_{i=1}^{n} \frac{(x+\alpha-1)!}{x!(\alpha-1)!}(0.5)^{x}(0.5)^{\alpha}$$

$$= \sum_{i=1}^{n}\left[\left(x_i+\alpha-1\right)! \bullet \frac{1}{(x_i)!} \bullet \frac{1}{\left[(\alpha-1)!\right]^{n}} \bullet 0.5^{\sum_{i=1}^{n} x_i} \bullet 0.5^{n\alpha} \right]$$

$$\ln L(\alpha) = \sum_{i=1}^{n} \ln(x_i+\alpha-1)! - \sum_{i=1}^{n} \ln x_i! - n\ln(\alpha-1)! + \sum_{i=1}^{n} x_i \ln 0.5 + n\alpha 0.5$$

$$\frac{d\ln L(\alpha)}{d\alpha} = \frac{d}{d\alpha}\left(\sum_{i=1}^{n} \ln(x_i+\alpha-1)! - n\ln(\alpha-1)! + n\alpha 0.5 \right)$$

$$n! \approx \sqrt{2\pi n} n^{n} e^{-n}$$

$$\therefore n! \approx \left(2\pi n\right)^{0.5} n^{n} e^{-n}$$

In this case therefore;

$$\sum_{i=1}^{n} \ln(x_i+\alpha-1)! = \sum_{i=1}^{n} \ln\left\{ \left[2\pi(x_i+\alpha-1)\right]^{0.5} \bullet (x_i+\alpha-1)^{x_i+\alpha-1} \bullet e^{-(x_i+\alpha-1)} \right\}$$

$$\frac{\partial}{\partial\alpha}\sum_{i=1}^{n} \ln(x_i+\alpha-1)! = \frac{\partial}{\partial\alpha}\sum_{i=1}^{n}\left[0.5\ln 2\pi + 0.5\ln(x_i+\alpha-1) + (x_i+\alpha-1)\ln\left(x_i+\alpha-1\right) - (x_i+\alpha-1) \right]$$

$$= \sum_{i=1}^{n}\left[\frac{0.5}{x_i+\alpha-1} \right] = \frac{0.5}{\sum_{i=1}^{n}(x_i+\alpha-1)} = \frac{0.5}{n(\bar{X}+\alpha-1)}$$

$$\ln(\alpha-1)! = \sum_{i=1}^{n} \ln\left\{ \left[2\pi(\alpha-1)\right]^{0.5} \bullet (\alpha-1)^{\alpha-1} \bullet e^{-(\alpha-1)} \right\}$$

$$\frac{\partial}{\partial\alpha}\ln(\alpha-1)! = \frac{\partial}{\partial\alpha}\sum_{i=1}^{n}\left[0.5\ln 2\pi + 0.5\ln(\alpha-1) + (\alpha-1)\ln\left(\alpha-1\right) - (\alpha-1) \right]$$

$$\frac{\partial}{\partial\alpha}\ln(\alpha-1)! = \frac{0.5}{\alpha-1}$$

Remember; $\dfrac{d\ln L(\alpha)}{d\alpha} = \dfrac{d}{d\alpha}\left(\sum_{i=1}^{n}\ln(x_i + \alpha - 1)! - n\ln(\alpha - 1)! + n\alpha 0.5\right)$

$$\frac{d\ln L(\alpha)}{d\alpha} = \frac{0.5}{n(\bar{X} + \alpha - 1)} - \frac{0.5n}{\alpha - 1} + n0.5$$

$$0 = \frac{0.5}{n(\bar{X} + \alpha - 1)} - \frac{0.5n}{\alpha - 1} + n0.5$$

which is nonlinear equation and can be found by numerical method.

We consider the Newton-Raphson method;

$$\hat{\alpha} = \alpha_0 - \frac{g(\alpha_0)}{g'(\alpha_0)}$$

Where;

$$g(\alpha_0) = \frac{\partial \ln L(\alpha)}{\partial \alpha}$$

$\alpha_0 =$ the initial estimate using the method of moments.

Therefore;

$$g(\alpha_0) = \left(\frac{0.5}{n(\bar{X} + \alpha - 1)} - \frac{0.5n}{\alpha - 1} + 0.5n\right)$$

$$g'(\alpha_0) = \frac{\partial}{\partial \alpha}\left(\frac{0.5}{n(\bar{X} + \alpha - 1)} - \frac{0.5n}{\alpha - 1} + n0.5\right) = \frac{0.5n}{(\alpha - 1)^2} - \frac{0.5}{n(\bar{X} + \alpha - 1)^2}$$

$\alpha_0 = \bar{X}$

4.4.3.2 The two parameter gamma mixing distribution

Let $f(x|\lambda)=\dfrac{e^{-\lambda}\lambda^{x}}{x!}$ and $g(\lambda)=\dfrac{\beta^{\alpha}}{\Gamma(\alpha)}e^{-\beta\lambda}\lambda^{\alpha-1}$

The mixture;

$$f(x)=\int_0^\infty \frac{e^{-\lambda}\lambda^{x}}{x!}\bullet\frac{\beta^{\alpha}}{\Gamma(\alpha)}e^{-\beta\lambda}\lambda^{\alpha-1}d\lambda$$

$$=\frac{\beta^{\alpha}}{x!\Gamma(\alpha)}\int_0^\infty e^{-\lambda(1+\beta)}\lambda^{x+\alpha-1}d\lambda$$

$$=\frac{\beta^{\alpha}}{x!\Gamma(\alpha)}\bullet\frac{\Gamma(x+\alpha)}{(1+\beta)^{x+\alpha}}$$

$$=\frac{\Gamma(x+\alpha)}{x!\Gamma(\alpha)}\bullet\frac{\beta^{\alpha}}{(1+\beta)^{x+\alpha}}$$

$$=\binom{x+\alpha-1}{x}\left(\frac{\beta}{1+\beta}\right)^{\alpha}\left(\frac{1}{1+\beta}\right)^{x};x=0,1,2,\cdots$$

which is a NBD with parameter α and $\dfrac{\beta}{1+\beta}$

In terms of pgf ;

$$H(S)=\int_0^\infty e^{\lambda(S-1)}g(\lambda)d\lambda$$

$$=\int_0^\infty e^{\lambda(S-1)}\frac{\beta^{\alpha}}{\Gamma(\alpha)}e^{-\beta\lambda}\lambda^{\alpha-1}d\lambda$$

$$=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\int_0^\infty e^{-\lambda[(1-S)+\beta]}\lambda^{\alpha-1}d\lambda$$

$$=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\bullet\frac{\Gamma(\alpha)}{(1-S+\beta)^{\alpha}}$$

$$=\frac{\beta^{\alpha}}{(1+\beta-S)^{\alpha}}$$

$$=\frac{\left(\dfrac{\beta}{1+\beta}\right)^{\alpha}}{\left(1-\dfrac{S}{1+\beta}\right)^{\alpha}}=\left[\frac{\dfrac{\beta}{1+\beta}}{1-\dfrac{S}{1+\beta}}\right]^{\alpha}$$

which is NBD with parameter α and $\dfrac{\beta}{1+\beta}$

The mean of the two parameter gamma mixing distribution

$$\bar{X} = H'(1)$$

$$H'(S) = \frac{\partial}{\partial S} H(S)$$

$$= \frac{\frac{\alpha}{1+\beta} \bullet \left(\frac{\beta}{1+\beta}\right)^{\alpha}}{\left[1 - \frac{S}{1+\beta}\right]^{\alpha+1}}$$

$$H'(1) = \frac{\frac{\alpha}{1+\beta} \bullet \left(\frac{\beta}{1+\beta}\right)^{\alpha}}{\left[1 - \frac{1}{1+\beta}\right]^{\alpha+1}}$$

$$= \frac{\alpha}{\beta}$$

$$H'(1) = \frac{\alpha}{\beta} = \bar{X} = Mean$$

The variance of the two parameter gamma mixing distribution

$$Var[X] = H''(1) + H'(1) - [H'(1)]^2$$

$$H''(S) = \frac{\partial}{\partial S} H'(S)$$

$$= \frac{\frac{(\alpha+1)\alpha\beta^{\alpha}}{(1+\beta)^{\alpha+2}}}{\left[1 - \frac{S}{1+\beta}\right]^{\alpha+2}}$$

$$H''(1) = \frac{(\alpha+1)\alpha\beta^{\alpha}}{(1+\beta)^{\alpha+2}} \bullet \frac{(1+\beta)^{\alpha+2}}{\beta^{\alpha}\beta^{2}}$$

$$Var[X] = \frac{\alpha(\alpha+1)}{\beta^2} + \frac{\alpha}{\beta} - \frac{\alpha^2}{\beta^2}$$

$$= \frac{\alpha + \alpha\beta}{\beta^2} = \frac{\alpha(1+\beta)}{\beta^2}$$

Estimation of the parameters of the two gamma mixing distribution

1. Method of moments

$$\bar{X} = \frac{\alpha}{\beta}$$

$$\alpha = \bar{X}\beta$$

$$S^2 = \frac{\alpha(1+\beta)}{\beta^2} = \frac{\bar{X}(1+\beta)}{\beta}$$

$$\hat{\beta} = \frac{\bar{X}}{S^2 - \bar{X}}$$

$$\hat{\alpha} = \bar{X}\beta = \bar{X}\left(\frac{\bar{X}}{S^2 - \bar{X}}\right) = \frac{\bar{X}^2}{S^2 - \bar{X}}$$

2. The maximum likelihood method

$$L = \prod_{i=1}^{n} \binom{x+\alpha-1}{x} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{1}{1+\beta}\right)^{x}$$

$$= \prod_{i=1}^{n} \frac{(x+\alpha-1)}{x!(\alpha-1)!} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{1}{1+\beta}\right)^{x}$$

$$= \prod_{i=1}^{n} \frac{\Gamma(x+\alpha)}{x!\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{1}{1+\beta}\right)^{x}$$

$$L = \sum_{i=1}^{n} \frac{\Gamma(x_i+\alpha)}{(x_i!)} \bullet \frac{1}{\left[\Gamma(\alpha)\right]^n} \bullet \left(\frac{\beta}{1+\beta}\right)^{n\alpha} \left(\frac{1}{1+\beta}\right)^{\sum_{i=1}^{n} x_i}$$

$$\ln L = \sum_{i=1}^{n} \ln\Gamma(x_i+\alpha) - \sum_{i=1}^{n} \ln(x_i!) - n\ln\Gamma(\alpha) + n\alpha\ln\left(\frac{\beta}{1+\beta}\right) + \sum_{i=1}^{n} x_i \ln\left(\frac{1}{1+\beta}\right)$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{\partial}{\partial \beta}\left\{ n\alpha[\ln\beta - \ln(1+\beta)] + \sum_{i=1}^{n} x_i[-\ln(1+\beta)] \right\}$$

$$= n\alpha\left(\frac{1}{\beta} - \frac{1}{1+\beta}\right) + \left(\sum_{i=1}^{n} x_i \bullet - \frac{1}{1+\beta}\right)$$

$$= \frac{n\alpha}{\beta(1+\beta)} - \frac{\sum_{i=1}^{n} x_i}{1+\beta}$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{n\alpha}{\beta(1+\beta)} - \frac{\sum_{i=1}^{n} x_i}{1+\beta}$$

$$0 = \frac{n\alpha}{\beta(1+\beta)} - \frac{\sum_{i=1}^{n} x_i}{1+\beta}$$

$$\frac{n\alpha}{\beta(1+\beta)} = \frac{\sum_{i=1}^{n} x_i}{1+\beta}$$

$$\frac{n\alpha}{\beta} = \sum_{i=1}^{n} x_i$$

$$\beta = \frac{n\alpha}{\sum_{i=1}^{n} x_i}$$

$$\breve{\beta} = \frac{\alpha}{\bar{X}}$$

$$\ln L = \sum_{i=1}^{n} \ln\Gamma(x_i + \alpha) - \sum_{i=1}^{n} \ln(x_i!) - n\ln\Gamma(\alpha) + n\alpha\ln\left(\frac{\beta}{1+\beta}\right) + \sum_{i=1}^{n} x_i \ln\left(\frac{1}{1+\beta}\right)$$

$$\frac{\partial \ln L}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left\{\sum_{i=1}^{n} \ln\Gamma(x_i + \alpha) - \sum_{i=1}^{n} \ln(x_i!) - n\ln\Gamma(\alpha) + n\alpha\ln\left(\frac{\beta}{1+\beta}\right) + \sum_{i=1}^{n} x_i \ln\left(\frac{1}{1+\beta}\right)\right\}$$

$$= \frac{\partial}{\partial \alpha}\left\{\sum_{i=1}^{n} \ln(x_i + \alpha - 1)! - \sum_{i=1}^{n} \ln(x_i!) - n\ln(\alpha - 1)! + n\alpha\ln\left(\frac{\beta}{1+\beta}\right) + \sum_{i=1}^{n} x_i \ln\left(\frac{1}{1+\beta}\right)\right\}$$

$$= \frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n\ln\left(\frac{\beta}{1+\beta}\right)$$

$$= \frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n\ln\beta - n\ln(1+\beta)$$

$$= \frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n\ln\frac{\alpha}{\bar{X}} - n\ln\left(1+\frac{\alpha}{\bar{X}}\right)$$

$$= \frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n\ln\alpha - n\ln\bar{X} - \ln(\bar{X}+\alpha) + n\ln\bar{X}$$

$$= \frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n\ln\alpha - \ln(\bar{X}+\alpha)$$

$$0 = \frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n\ln\alpha - \ln(\bar{X}+\alpha)$$

which is a non-linear equation and needs to be solved by numerical method, using the Newtonôs method:

$$\alpha_0 = \frac{\bar{X}^2}{S^2 - \bar{X}}$$

$$g(\alpha_0) = \frac{\partial \ln L(\alpha)}{\partial \alpha} = \frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n\ln\alpha - \ln(\bar{X}+\alpha)$$

$$g'(\alpha_0) = \frac{\partial g(\alpha_0)}{\partial \alpha} = \frac{0.5n}{(\alpha-1)^2} - \frac{0.5}{n(\bar{X}+\alpha-1)} + \frac{n}{\alpha} - \frac{1}{(\bar{X}+\alpha)}$$

$$\alpha = \alpha_0 - \frac{g(\alpha_0)}{g'(\alpha_0)}$$

4.5 <u>THE POISSON LINDLEY DISTRIBUTION</u>

Assuming that $f(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$ and $g(\lambda) = \frac{\theta^2}{\theta+1}(\lambda+1)e^{-\lambda\theta}; \lambda, \theta > 0$

Then, $f(x) = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!} \bullet \frac{\theta^2}{\theta+1}(\lambda+1)e^{-\lambda\theta}\, d\lambda$

$$= \frac{\theta^2}{x!(\theta+1)} \int_0^\infty e^{-\lambda(1+\theta)} (\lambda^{x+1} + \lambda^x) d\lambda$$

$$= \frac{\theta^2}{x!(\theta+1)} \left\{ \int_0^\infty e^{-\lambda(1+\theta)} \lambda^{x+1} d\lambda + \int_0^\infty e^{-\lambda(1+\theta)} \lambda^x d\lambda \right\}$$

$$= \frac{\theta^2}{x!(\theta+1)} \left\{ \frac{\Gamma(x+2)}{(1+\theta)^{x+2}} + \frac{\Gamma(x+1)}{(1+\theta)^{x+1}} \right\} property 1.3$$

$$= \frac{\theta^2}{(\theta+1)} \left\{ \frac{x+1}{(1+\theta)^{x+2}} + \frac{1}{(1+\theta)^{x+1}} \right\}$$

$$= \frac{\theta^2}{(\theta+1)^2} \left\{ \frac{x+1}{(1+\theta)^{x+1}} + \frac{1}{(1+\theta)^{x}} \right\}$$

$$= \frac{\theta^2}{(\theta+1)^{2+x}} \left\{ \frac{x+1}{(1+\theta)} + 1 \right\} = \frac{\theta^2}{(\theta+1)^{2+x}} \left\{ \frac{x+2+\theta}{1+\theta} \right\}$$

$$f(x) = \frac{\theta^2 (x+2+\theta)}{(\theta+1)^{3+x}}$$

In terms of pgf;

$$H(S) = \int_0^\infty e^{-(-S+1)} g(\lambda) d\lambda$$

$$= \int_0^\infty e^{-\lambda(1-S)} \frac{\theta^2}{\theta+1} (\lambda+1) e^{-\lambda\theta} d\lambda$$

$$= \frac{\theta^2}{\theta+1} \int_0^\infty e^{-\lambda[(1-S)+\theta]} \bullet (\lambda+1) d\lambda$$

$$= \frac{\theta^2}{\theta+1} \int_0^\infty e^{-\lambda[(1-S)+\theta]} \lambda d\lambda + \int_0^\infty e^{-\lambda[(1-S)+\theta]} \lambda^0 d\lambda$$

$$= \frac{\theta^2}{\theta+1} \left[\frac{\Gamma(2)}{(-S+1+\theta)^2} + \frac{\Gamma(1)}{(-S+1+\theta)} \right]$$

$$= \frac{\theta^2}{\theta+1} \left[\frac{\Gamma(1) \bullet 1}{(-S+1+\theta)^2} + \frac{\Gamma(1)}{(-S+1+\theta)} \right]$$

$$= \frac{\theta^2}{\theta+1} \left[\frac{1}{(-S+1+\theta)^2} + \frac{1}{(-S+1+\theta)} \right] \mathrm{Pr} operty 1.2$$

$$H(S) = \frac{\theta^2}{\theta+1} \bullet \frac{2-S+\theta}{(1+\theta-S)^2}$$

The mean of the Poisson Lindley distribution

$$E[X] = H'(1)$$

$$H'(S) = \frac{d}{dS} H(S)$$

$$= \frac{d}{dS}\left[\frac{\theta}{\theta+1} \bullet \frac{2-S+\theta}{(1+\theta-S)^2} \right]$$

$$= \frac{\theta^3 + 3\theta^2 - S\theta^2}{(\theta+1)(\theta+1-S)^3}$$

$$H'(1) = \frac{\theta^3 + 3\theta^2 - S\theta^2}{(\theta+1)\theta^3}$$

$$= \frac{\theta^3 + 2\theta^2}{\theta^3(\theta+1)}$$

$$= \frac{\theta^2(\theta+2)}{\theta^3(\theta+1)}$$

$$H'(1) = \frac{(\theta+2)}{\theta(\theta+1)} = \overline{X} = Mean$$

The variance of the Poisson Lindley distribution

$$Var[X] = H''(1) + H'(1) - [H'(1)]^2$$

$$H''(S) = \frac{d}{dS} H'(S)$$

$$= \frac{d}{dS} \frac{\theta^3 + 3\theta^2 - S\theta^2}{(\theta+1)(\theta+1-S)^3}$$

$$= \frac{-\theta^2}{(\theta+1)(\theta+1-S)^3} + \frac{3\left(\theta^3 + 3\theta^2 - S\theta^2\right)}{(\theta+1)(\theta+1-S)^4}$$

$$= \frac{2\theta^3 + 8\theta^2 - 2\theta^2}{(\theta+1)(\theta+1-S)^4}$$

$$H''(1)=\frac{2\theta^3+8\theta^2-2\theta^2}{(\theta+1)\theta^4}$$

$$Var[X]=\frac{2\theta^3+8\theta^2-2\theta^2}{(\theta+1)\theta^4}+\frac{(\theta+2)}{\theta(\theta+1)}-\frac{(\theta+2)^2}{\theta^2(\theta+1)^2}$$

$$Var[X]=\frac{4\theta^4+6\theta^3+2\theta^2+\theta^5}{\theta^4(\theta+1)^2}=\frac{\theta^2(4\theta^2+6\theta+2+\theta^5)}{\theta^4(\theta+1)^2}$$

Estimation of the parameter θ in the Poisson Lindley distribution

1) **The method of moments.**

$$E[X]=\bar{X}=\frac{\theta+2}{\theta(\theta+1)}$$

$$\bar{X}=\frac{\theta+2}{\theta^2+\theta}$$

$$\bar{X}\theta^2+\bar{X}\theta=\theta+2$$

$$\bar{X}\theta^2+\bar{X}\theta-\theta-2=0$$

$$\bar{X}\theta^2+\theta(\bar{X}-1)-2=0$$

$$\therefore\ \theta=\frac{-(\bar{X}-1)+\sqrt{(\bar{X}-1)^2+8\bar{X}}}{2\bar{X}}$$

2) **The maximum likelihood method.**

$$L=\prod_{i=1}^{n}\frac{\theta^2(x+2+\theta)}{(\theta+1)^{3+x}}$$

$$L=\frac{\theta^{2n}\sum_{i=1}^{n}\left(x_i+2+\theta\right)}{\left(\theta+1\right)^{3n}\left(\theta+1\right)^{\sum_{i=1}^{n}x_i}}$$

$$\ln L = 2n\ln\theta + \ln\sum_{i=1}^{n}(x_i + 2 + \theta) - 3n\ln(\theta+1) - \sum_{i=1}^{n}x_i \ln(\theta+1)$$

$$\frac{d\ln L}{d\theta} = \frac{2n}{\theta} + \frac{1}{\sum_{i=1}^{n}(x_i + 2 + \theta)} - \frac{3n}{(\theta+1)} - \frac{\sum_{i=1}^{n}x_i}{\theta+1}$$

$$0 = \frac{2n}{\theta} + \frac{1}{\sum_{i=1}^{n}(x_i + 2 + \theta)} - \frac{3n}{(\theta+1)} - \frac{\sum_{i=1}^{n}x_i}{\theta+1}$$

which is nonlinear and needs to be solved by numerical method, like the Newtonôs method.

CHAPTER V

DATA ANALYSIS

5.1 INTRODUCTION

In this chapter, fitting data is done using the Geometric, NBD and the Poisson Lindley distributions.

Secondary data used is by Luc Tremblay (1992), and the analysis is based on excel.

No of claims Per policy (x)	Frequency No of policies (f)	x*f	$(x-\bar{x})$	$(x-\bar{x})^2$	$f \times (x-\bar{x})^2$
0	103,704	0	-0.15514	0.024068	2,495.99139
1	14,075	14,075	0.84486	0.713788	10,046.57090
2	1,766	3,532	1.84486	3.403508	6,010.59557
3	255	765	2.84486	8.093228	2,063.77318
4	45	180	3.84486	14.782948	665.23266
5	6	30	4.84486	23.472668	140.83601
6	2	12	5.84486	34.162388	68.32478
Total	119,853	18,594			21,491.3245

Mean $= \bar{x} = \dfrac{18594}{119853} = 0.15514$

Variance $= \dfrac{21491.3245}{119853} = 0.17931403$

5.1.1 <u>Method of moments</u>

- For Geometric distribution;

$$\breve{\mu} = \frac{1}{\bar{X}}$$

$$\breve{\mu} = \frac{1}{0.15514} = 6.44579$$

- For the one parameter NBD;

$$\hat{\alpha} = \bar{X}$$

$$\hat{\alpha} = 0.15514$$

- For the two parameter NBD;

$$\hat{\alpha} = \frac{\bar{X}}{S^2 - \bar{X}}$$

$$\hat{\alpha} = \frac{0.15514^2}{0.17931 - 0.15514} = 0.99563$$

$$\hat{\beta} = \frac{\bar{X}}{S^2 - \bar{X}}$$

$$\hat{\beta} = \frac{0.15514}{0.17931 - 0.15514} = 6.41765$$

- For Poisson Lindley;

$$\hat{\theta} = \frac{-(\bar{X}-1) + \sqrt{(\bar{X}-1)^2 + 8\bar{X}}}{2\bar{X}}$$

$$\hat{\theta} = \frac{-(0.15514-1) + \sqrt{(0.15514-1)^2 + (8*0.15514)}}{2 \times 0.15514} = 7.22908$$

Method of moments

No of claims (x)	Observed data	Poisson distribution	Geometric distribution	One parameter gamma	Two parameter gamma	Poisson Lindley distribution
0	103, 704	102,629.559	103,756.255	107,633.414	19,294.257	103,733.629
1	14,075	15,921.950	13,934.887	8,349.124	32,326.655	13,971.601
2	1,766	1,235.066	1,871.512	2,411.102	13,534.686	1,863.813
3	255	63.869	251.352	866.044	11,356.665	246.661
4	45	2.477	33.758	341.561	9,532.608	32.425
5	6	0.07686	4.5338	141.923	8,003.274	4.238
6	2	0.00199	0.6089	60.970	6,720.268	0.551
Total	119,853	119,853	119,852.906	119,804.138	100,768.413	119,852.918

5.1.2 Maximum likelihood method

- For Geometric distribution;

$$\hat{p} = \frac{1}{\bar{X}}$$

$$\hat{p} = \frac{1}{0.15514} = 6.44579$$

- For the one parameter NBD;

$$\hat{\alpha} = \alpha_0 - \frac{g(\alpha_0)}{g'(\alpha_0)}$$

$$\alpha_0 = \bar{X} = 0.15514$$

$$g(\alpha_0) = \left(\frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + 0.5n\right) = \left(\frac{0.5}{119853(0.15514+0.15514-1} - \frac{0.5\times119853}{0.15514-1} + (0.5\times119853)\right)$$

$$= 130857.2$$

$$g'(\alpha_0) = \frac{\partial}{\partial\alpha}\left(\frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n0.5\right) = \frac{0.5n}{(\alpha-1)^2} - \frac{0.5}{n(\bar{X}+\alpha-1)^2}$$

$$= \frac{0.5\times 119853}{(0.15514-1)^2} - \frac{0.5}{119853(0.15514+0.15514-1)^2} = 83955.55$$

$$\hat{\alpha} = 0.15514 - \frac{130857.2}{83955.55} = -1.40351$$

- For the two parameter NBD;

$$\alpha_0 = \frac{\bar{X}^2}{S^2 - \bar{X}} = 0.99563248$$

$$g(\alpha_0) = \frac{\partial \ln L(\alpha)}{\partial\alpha} = \frac{0.5}{n(\bar{X}+\alpha-1)} - \frac{0.5n}{\alpha-1} + n\ln\alpha - \ln(\bar{X}+\alpha)$$

$$= \frac{0.5}{119853(0.15514+0.9956-1)} - \frac{0.5\times 119853}{0.9956-1} + (119853\times\ln 0.9956) - \ln(0.15514+0.9956)$$

$$= 13720423.6$$

$$g'(\alpha_0) = \frac{\partial g(\alpha_0)}{\partial\alpha} = \frac{0.5n}{(\alpha-1)^2} - \frac{0.5}{n(\bar{X}+\alpha-1)} + \frac{n}{\alpha} - \frac{1}{(\bar{X}+\alpha)}$$

$$= \frac{0.5\times 119853}{(0.9956-1)^2} - \frac{0.5}{119853(0.15514+0.9956-1)} + \frac{119853}{0.9956} - \frac{1}{(0.15514+0.9956)}$$

$$= 3141709200$$

$$\hat{\alpha} = \alpha_0 - \frac{g(\alpha_0)}{g'(\alpha_0)}$$

$$= 0.9956 - \frac{13720423.6}{3141709200} = 0.9913$$

$$\hat{\beta} = \frac{\alpha}{\bar{X}} = \frac{0.99126}{0.15514} = 6.3895$$

- For Poisson Lindley distribution

$$\hat{\theta} = \theta_0 - \frac{g(\theta_0)}{g'(\theta_0)}$$

$$\theta_0 = 7.2291$$

$$g(\theta_0) = \frac{2n}{\theta} + \frac{1}{\sum_{i=1}^{n}(x_i + 2 + \theta)} - \frac{3n}{(\theta+1)} - \frac{\sum_{i=1}^{n} x_i}{\theta+1} = \frac{2n}{\theta} + \frac{1}{n\overline{X} + 2n + n\theta} - \frac{3n}{(\theta+1)} - \frac{n\overline{X}}{\theta+1}$$

$$= \frac{2\times 119853}{7.2291} + \frac{1}{(119853\times 0.15514) + (2\times 119853) + (119853\times 7.2291)} - \frac{3\times 119853}{(7.2291+1)} - \frac{119853\times 0.155}{(7.2291+1)}$$

$$= -12794.7$$

$$g'(\theta_0) = \frac{3n}{(\theta+1)^2} + \frac{n\overline{X}}{(\theta+1)^2} - \frac{2n}{\theta^2} - \frac{1}{(n\overline{X} + 2n + n\theta)}$$

$$= \frac{3\times 119853}{(7.2291+1)^2} + \frac{119853\times 0.15514}{(7.2291+1)^2} - \frac{2\times 119853}{7.2291^2} - \frac{1}{[(119853\times 0.15514) + (2\times 119853) + 7.2291]^2}$$

$$= 20.05686$$

Maximum Likelihood Method

No of claims (x)	Observed data	Poisson distribution	Geometric distribution	NBD based on one parameter gamma	NBD based on two parameters gamma	Poisson Lindley distribution
0	103, 704	102,629.559	103,756.255	317,064.079	103,765.315	113,903.654
1	14,075	15,921.950	13,934.887	(222,501.329)	13,919.629	5,654.581
2	1,766	1,235.066	1,871.512	22,445.325	1,875.480	280.185
3	255	63.869	251.352	2,231.408	253.065	13.859
4	45	2.477	33.758	445.303	34.172	0.6845
5	6	0.07686	4.5338	115.623	4.6163	0.03375
6	2	0.00199	0.6089	34.653	0.6238	0.001662
Total	119,853	119,853	119,852.906	119,835.062	119,852.901	119,853

5.2 REMARKS AND RECOMMENDATIONS

From the method of momentŝ table, the Geometric and Poisson Lindley distributions compare well with the observed data (seems to give better results) while the Geometric and two parameter gamma distributions compare well with the observed data, from the maximum likelihood table. The Geometric distribution is clearly a better fit in both methods.

The data used may not give a true representation of the claims, as secondary data was used. A current claims experience will be a better data for the same research, and other mixtures other than the Poisson can be looked into to establish if they are better than the ones looked at.

REFERENCES

Das and Basu. (2003). A Markov Chain Based No-Claim-Discount System.

The Institute and Faculty of Actuaries Examination, CT 4. (2009, 2012).

Lemaire and Zi. (1994). A comparative analysis of 30 Bonus-Malus Systems. *ASTIN BULLETIN, Vol. 24, No.2, 1994.*

Emarmverdi, Firouzi and Emamdi. (2013). Design of an optimal Bonus-Malus System for automobile insurance &moving between two different scales.

Nath and Sinha. (2014). A Markovian study of no claim discount system of Insurance Regulatory and Development Authority and its application.

Ade Ibiwoye, I. A. Adekele and S. A. Aduloju. (2011). Quest for Optimal Bonus-Malus in Automobile Insurance in Developing Economies: An Actuarial Perspective.

Soren Asmussen. (2014). Modelling and Performance of Bonus-Malus Systems: Stationarity versus Age-Correction.

Jean Lemaire. (1998). Bonus-Malus Systems: The European and Asian approach to Merit-rating.

J. F. Walhin and J. Paris. (1999). Using mixed Poisson processes in connection with Bonus-Malus Systems. ASTIN BULLETIN, Vol. 29, pp 81-99.

Laurence Hulin and Daniel Justens. General properties of mixed Poisson distributions Application to data fitting in case of ñautomobile collision claimsò distributions.

Alexander Volfovsky. (2007). Markov Chains and Applications.

Michel Denuit and Jan Dhaene. (2000). Bonus-Malus Scales using Exponential loss functions.

Ales Kozubik. (2006). Bonus-Malus systems as Markov process and hungry for bonus. Journal of Information, Control and Management Systems, Vol. 4, (2006), No. 1.

Niemiec Malgorzata. (2007). A Bonus-Malus System as a Markov Set-Chain.

Guy H. Whitehead. (1991). No-Claim Discount or Bonus/Malus systems in Europe.

Park,Lemaire and Chua. (2010). Is the Design of Bonus-Malus Systems Influenced by Insurance Maturity or National Culture? ï Evidence from Asia.

Miklos A. and Laszlo M. (2012). Estimation of Claims Numbers in Automobile Insurance.

Philip J. Boland. (2006). Statistical Methods in General Insurance.

Nicholas E. Frangos and Sypridon D. Vrontos. (2001). Design of Optimal Bonus-Malus Systems with a Frequency and Severity Component on an individual basis in automobile insurance. ASTIN BULLETIN, Vol. 31, No. 1, pp 1-22.

Fabio Baione,Susanna Levantesi and Massimiliano Menzietti. (2002). The Development of an Optimal Bonus-Malus System in a Competitive Market. ASTIN BULLETIN, Vol. 32, No. 1, pp 159 ï 170.

Jon Holtan. (2001). Optimal Loss Financing Under Bonus-Malus Contracts. ASTIN BULLETIN, Vol. 31, No. 1, pp 161 ï 173.

Jean Pinquet. (2012). Experience Rating in Non-Life Insurance.