University of Nairobi

School of Mathematics

ISSN: 2410-1397

Master Project in Actuarial Science

Designing of Optimal Bonus-Malus Systems Based on Individual Characteristics in Automobile Insurance

Research Report in Actuarial Science, Number 36, 2019

Koech Jairus Kipkorir

June 2019

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

Master Project in Actuarial Science **University of Nairobi**
June 2019

Designing of Optimal Bonus-Malus Systems Based on Individual Characteristics in Automobile Insurance

Research Report in Actuarial Science, Number 36, 2019

Koech Jairus Kipkorir

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

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

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

Abstract

Actuarial Science is the discipline that concerns with uncertain events where the concept of probability and statistics provide for essential instruments in the measurement and management of risks in insurance and finance.A key aspect of the business of insurance is the calculation of the price to pay commonly known as the premium to pay in exchange for the transfer of risk.Many insurance companies charge premiums on the policyholders based only on the claim frequency.That way a policyholder who underwent an accident with a small size of loss will be unfairly punished in comparison to an insured policyholder who had an accident with a large amount of loss.In automobile liability insurance,the policyholders do not all the same risk to have an accident.The premium that is charged to each policyholder has to be proportionate to his/her underlying risk to have an accident.Motivated by this,we consider the design of a model that incorporate both the frequency and and severity components and we suggest a method that deliberate concurrently on the number of claims,the exact size of loss and the individual characteristics.
The modeling of claim frequency component is based on Poisson mixtures where the number of claims is distributed according to the negative binomial type I.The severity component is modeled using the exponential mixtures where the the losses are distributed according to a Pareto distribution.Using the Baye's theorem we get the posterior function for the number of claims and the claim amount component.Considering only the claim frequency the premium was estimated as the mean of the posterior structure function in computing premiums.The premiums based on both frequency and severity component was estimated as the product of the mean of posterior structure function of the frequency and severity component.

Keywords
Optimal BMS,claim frequency,claim severity,poisson mixtures,exponential mixtures,posterior structure function,a priori classification criteria, a posteriori classification criteria,

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

DISTRIBUTOR: School of Mathematics, University of Nairobi, Kenya

Declaration and Approval

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

Signature Date

KOECH JAIRUS KIPKORIR
Reg No. I56/7003/2017

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

Signature Date

Prof. Patrick Weke
School of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: pweke@uonbi.ac.ke

Signature Date

Prof. J.A.M Ottieno
School of Mathematics
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: jottieno@uonbi.ac.ke

Dedication

This project is dedicated to my parents and my beloved family.

Contents

Acknowledgments

First, I thank the Almighty God for giving me the wisdom,knowledge and strength to pursue and complete this project.Secondly, i am very grateful for my supervisor Prof. J.A.M Ottieno,for giving me the chance to work in this challenging subject of Bonus-Malus system and for guiding and correcting brilliantly.He is a wonderful advisor,not only knowledgeable but truly cares for his students.I am also very grateful to Prof. Weke for his assistance and valuable suggestions for my project and also for the help he has offered me over the past two years.I would have never finished my project without them.Finally,i would like to thank my classmates and the entire staff members of School of Mathematics for supporting me in one way or another.....

Koech Jairus Kipkorir

Nairobi, 2017.

1 General Introduction

1.1 Background/Historical Information

The basic concept of insurance consist of creating a portfolio where the risks of the insured are managed.Since the risks in the portfolio are not the same for every insured ,the premium paid by each member should be equivalent to risk that an insured exposes the portfolio.In insurance companies that deal with third-party liability insurance especially the automobile,the risk structure of the portfolio of policyholders is heterogeneous in that the risk are not equal, i.e. the policyholders do not all have the same risk to have an accident.The detection and measurement of the elements that affect and transform the risk of an accident and hence the premium that must be paid,is essential for the designing fair tariff that fairy allocates the cost of claims to each and every insured individual.

The main duty of an actuary in modelling of a new pricing structure is to ensure that it equitable .This is achieved with the partition of the portfolio of policies into similar groups and all the policyholders belonging to the same group paying equal premium.The more heterogeneity of risk structure exist inside each class of policyholders the more unfair the tariff classification method is as policyholders with a different probability to have an accident pay the same premium.This partition can be done using information known to the insurer before the insured join the portfolio or a posteriori information or by using both of them.The most frequently used a priori classification criteria are the automobile type and use, age and sex of the insured,cubic capacity of the engine and place where the insured resides.After the use of a priori classification criteria the tariff structure classes are still homogeneous ,hence its is suggested taking into consideration all these differences in the posteriori,by modifying the premium according to the each policyholder's claim history.

According to the philosophy of a posteriori classification criteria the driver's past claim behavior give the best forecast of the future claim numbers.The evolution of the philosophy of the a posteriori classification has guided the growth of the classification systems which are known in almost every country around the world Bonus-Malus Systems (BMS).The Bonus-Malus-Systems penalize policyholders who caused and and made a claim by premium increase or and appreciate the policyholders who did not caused and accident during the period under consideration by awarding the discounts of the premium of bonuses.The main purpose of a BMS is to enable the insurance companies under how they charge premiums to their customers proportionately based on the characteristics of each insured person at any given year.Such characteristics includes the number of claims,the amount of each claims and driving abilities

The optimal BMS obtained using the frequency component has the disadvantage of penalizing the policyholders without taking into consideration the amount of claim that the claim caused.As a matter of fact,all Bonus-Malus Systems considered do not factor in the amount of loss on renewal of the policy and as such ,the policyholders who made a claim with small loss are forced to renew their policy with the same price with the policyholders which the insurance company pay high cost in paying their claims.In this sense a BMS which can separate the policyholders according to the number of claims and amounts of their claims should be developed.

Besides,the current Bonus-Malus System does not take into consideration the type of accident and penalize equally the accidents which cause a property damage claim only and the claims which cause property damage and bodily injury claims.The bodily injury claims are of great importance because even though they represent a small percent of number of claims,they cost a serious percent of the total claim amounts.

1.2 Definitions,Notations and Terminologies

Mixtures

Mixtures arises when a probability density function $f(x/\nu)$ depends on a parameter ν that is uncertain and is itself a random variable with density $g(\nu)$.Then taking the weighted average of $f(x/\nu)$ with $g(\nu)$ as weight produces the mixture distribution.
The pdf of a continuous mixture is given by

$f(x) = \int_0^\infty f(x/\nu)g(\nu)d\nu$

where,

$f(x/\nu) =$ the conditional pdf

and,

$g(\nu) =$ the continuous mixing distribution.

pdf= probability density function

BMS=Bonus-malus system

NCD= No claim discount system

PMF =Probability Mass Function

$E(x^r)$= The rth Moment of the mixture

1.3 Research Problem

Originally Bonus-Malus System was obtained by considering claim frequency without taking into account the size of claim.This system was unfair since it punishes the claim numbers independently of their severity,that is without taking into account the size of loss.
The early studies considered Poisson distribution in obtaining claim frequency component which assumes homogeneity in policyholders risks but in reality different individuals have different underlying risk characteristics and hence the need to take into consideration heterogeneity.Equidispersion is a major disadvantage of Poisson distribution where variance is to be equal to the mean which may not be consistent with observed data.The development of specific models (mixture models) to represent different characteristics of data was motivated by high presence of overdispersion in the data.

The study further proposed a BMS that incorporate both information known to the insurer and information that the insured exhibits during the period of observation and basing on each policyholder's characteristics.

1.4 Objectives

To obtain optimal Bonus Malus system according to the claim numbers ,claims size component and on characteristics of each individual policyholder.

Specific Objectives

The specific objectives are to :

1.To estimate number of claims and claim size component according to posteriori criteria.

2.To estimate number of claims and claim size component according to a priori and posteriori criteria.

3.To compare premiums for various claim frequency with claim severity during the first period of observation.

4.Estimates premiums using generalized models based on information known to the insurer before the insured join the portfolio and after taking into account the characteristics of each insured.

1.5 Methodology

The following are methods that are used to achieve the above objectives.

Continuous Mixtures

Poisson Mixtures

Let

$f(x/\nu) = \frac{\exp(\nu)\nu^x}{x!}$ $\nu > 0$ and $x = 0, 1, 2, ...$

Then

$f(x) = \int_0^\infty \frac{\exp(\nu)\nu^x}{x!} g(\nu) d\nu$

$E(x) = \nu$ and $Var(x) = \nu$

Exponential Mixtures

We consider exponential distribution of type II whose mean is the parameter ν

$f(x/\nu) = \frac{1}{\nu}\exp(-\frac{x}{\nu})$ $\nu > 0$ $x > 0$

$f(x) = \int_0^\infty \frac{1}{\nu}\exp(-\frac{x}{\nu}) g(\nu)$

$E(x) = \frac{1}{v}$ and $Var(x) = \frac{1}{v^2}$

The Gamma function

$g(v) = \frac{\beta^{\alpha} v^{\alpha-a} \exp(-\beta v)}{\Gamma(\alpha)}, v > 0, \alpha > 0, \beta > 0$

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

$\int_0^{\infty} \exp(-\beta t) t^{\alpha-1} dt = \frac{\Gamma(\alpha)}{\beta^{\alpha}}$

Inverse Gamma Function

$$g(y) = \frac{\frac{1}{m}\exp(-\frac{m}{y})}{(\frac{y}{m})^{s+1}\Gamma(\alpha)}$$

$E(y) = \frac{m}{s-1}$ and $Var(y) = \frac{m^2}{(s-1)^2(s-2)}$

Lindley Function

$g(y) = \frac{\vartheta^2}{\vartheta+1}(y+1)\exp(ya\vartheta)$, $y > 0$ and $\vartheta > 0$

$E(y) = \frac{\vartheta^2+4\vartheta+2}{\vartheta^2(\vartheta 01)^2}$

Estimation

The parameters of the above distributions are estmated using the following methods;

1) Method of Moments

In this method,the parameters by method of moments we equate the representative variance s^2 and representative mean $\bar{x}$ to the corresponding population values.

2)Maximum likelihood Estimate method

The method of maximum likelihood estimate is based on the likelihood function,$L(\vartheta;x)$.Suppose we are given a statistical model $\{f(\cdot;\vartheta)/\vartheta \in \vartheta\}$, where ϑ denotes the parameter for the distribution.The method of maximum likelihood estimate find the values of the model parameter ,ϑ, that maximize the likelihood function ,$L(\vartheta;x)$.
This estimator is given by,

$$l(\vartheta;x) = lnL(\vartheta;x)$$

3) Bayes Theorem-Bayesian Inference

Consider observations $n_1, n_2, ..., n_t$,and suppose we wan to estimate the parameters $v_1, v_2, ..., v_n$.Let us denote $v = (v_1, v_2, ..., v_n)^T$ the vector of the parameters and $n = (n_1, n_2, ..., n_t)^T$ observed data with likelihood function $L(k/v)$.Suppose $u(v)$ is the prior distribution for v, which denotes our subjective belief or the prior information we have about v.The posterior distribution of the vector of parameters v will be obtained using the Bayes theorem,and when the parameter v is continuous valued,which is most common situation,will be the following:

$$f(v/n) = \frac{L(n/v) * g(v)}{\int_0^\infty L(n/v) * g(v) dv}$$

1.6 Literature Review

Generally in the insurance industry,a BMS are experience rating mechanism which punishes policyholders that makes one or more accidents by premium surcharge and appreciate by giving discounts to policyholders who had no claim in any period under consideration.BMS are common in automobile insurance industry.The BMS can also be called a No - Claim Discount (NCD) or no - claim bonus in Britain and Australia.

The basic concept of NCD is that higher insurance costs that are charged on average to every policyholder that corresponds to the high number of the claims.

G. Dionne and C. Vanasse (1989,1992) Builds a Bonus -Malus system in which a priori and a posteriori information are integrated with individual characteristics so as to modify premiums of individual policyholder with a given period of time .He develops a statistical model that integrates sufficiently categories of risks and experience rating .He used Poisson and Negative binomial models with a regression component in order to use all the information available to estimate accidents distribution,.The parameters of negative binomial regression model were estimated using maximum likelihood estimate.The premiums were calculated using the expected value principle and insurance table were obtained as a function of time,past number of accidents and remarkable variables in the regression.

Luc Tremblay (1992) proposed a bonus malus system by fitting of data using the poisson inverse Gaussian Distribution extending the model introduced by Lemaire (1976) thus minimizing the average total risk of insurer since the insurer is at risk.This model is based on the number of claims N which is random irrespective of their amount.He represent claim frequency Poisson distribution with mean v which is a random variable with distribution representing the expected risk inherent in any given portfolio.He assumed the the distribution of v is inverse Gaussian since it has thick tails and provide an advantage of having a closed form expression for the moment generating function.The Mixed poisson provide a better fit from the insurer's perspective because the variance exceeds its mean.To minimize the insurer's risk he estimated the posterior distribution of v using the Bayesian theorem and estimated the parameters using the maximum likelihood estimate.He used the principle of zero utility in order to determine the premium .

Nikolaous Frangos and Dimitris (2004) Proposed a model for modelling losses using exponential-Inverse Gaussian distribution allowing for covariates.The model is preferred to Pareto distribution because it has a shorter tail and considered appropriate for modelling data without larger tails.The claim losses are distributrd according to exponential distribution with mean y.Since the policyholders do not have the same mean for the claim amounts , they expressed inform of a distribution known as inverse Gaussian distribution.The mixed exponential distribution provide a good fit for claim size data with small tails.The allowance of covariates (regression coefficient) in the model enables the modelling of data with different characteristics of policyholders.The model parameters were estimated using the maximum likelihood estimation through the EM algorithm.The posterior expectation was estimated using Bayes theorem making use of modified Bessel Function.The covariates were fitted using exponential General Linear Model via EM algorithm.

Spyridon D. Vrontos et al (1998) Introduced a BMS where past number of accidents and the correct claim amount for every accident caused are considered.In particular, the

BMS suggested allocates to every individual policyholder a premium equivalent to past number of accidents they have and the exact amount of loss that the claim caused.That is ,the bigger the claim size the bigger the premium that the policyholder has to pay.He considered the frequency and severity to be independent in order to able to deal with each component separately.He used the Negative Binomial distribution to represent the number of claims and Pareto distribution for the amount of loss caused and the net premium calculation principle.

Weihang Ni,Carina Contantinescu and Athanasios A. Pantelous (2014) developed a bonus malus systems with claims severities distributed according to a Weibull distribution which addresses the bonus hunger problem.The modelling of claims is done by mixing Poisson intensity v with gamma which give rise to Negative Binomial distribution which addresses overdispersion in the data.In addition,they applied the Bayesian theorem to obtain the posterior distribution and subsequently the posterior mean. They also assumed that the claim amount is distributed according to exponential type I distribution with mean $\frac{1}{\vartheta}$.The Levy ($\frac{1}{2}$) stable)is used to describe because they are not equal for all the insured .The mixture of exponential type I with Levy distribution result in a Weibull distribution which does not have long tails as Pareto which reinsurance companies rely on to alleviate the burden of extremely large claims.The Bayes' theorem is used to obtain the posterior distribution and subsequently the posterior mean.The premiums are obtained using the net premium principle .The application of Weibull distribution to claim size data shows that the initial premiums payment are lower than the Pareto distribution which is more preferred by a starting policyholder which creates more competitiveness to the insurer.

Emad Abdelgalil Ali Ismail (2016) present the design of an optimal BMS according to finite mixture models with claim sizes distributed according to Gamma.This BMS is designed using the Bayesian approach,probability distribution and taking into account the Poisson distribution to represent number of claims.The number of claims was assumed to be Poisson distribution and the basic risk was assumed to b Gamma resulting in Negative binomial distribution. Using the Bayesian approach he obtained the posterior structure function and posterior mean.The claim sizes was modelled sing Gamma distribution as the claim size and the prior function as the Gamma function resulting in Gamma distribution with updated parameters.The parameters were estimated using the maximum likelihood estimate method.The premiums obtained using the net premium principle as the product of posterior mean based on claim numbers and amount of claim.The model overcomes the limitations of the model proposed by Weihong according to Weibull distribution and negative distribution where premiums increases with increase in total number of claims to a given limit then premiums reduces with increase in claim numbers.

George Tzougas et al (2017) presented the design of optimal BMS using various finite mixtures of distribution and regression models.He extended the actuarial literature research which uses generalized linear models for pricing of risks through ratemaking based on a priori risk classification (see,Denuit et al 2007) by considering GMLSS.For the frequency component,the number of claims is assumed to have the following distribution; a finite Poisson,Delaporte and Negative Binomial Mixture, and for the severity component they considered that the losses the following distribution; a finite exponential,Gamma and Weibull Mixture.He expanded Lemaire (1995) set up and applied baye's theorem to obtain the posterior probability of the policyholders.

George Tzougas et al (2018) presents the design of BMS with two component mixture models emerging from no-identical parametric families.They made the following contribution to the present study:a)Developed a new way of designing optimal bonus-malus system considering both claim numbers and amount without assuming that the distribution for claim numbers and amounts come from the same parametric family.b)Developed a bonus malus system by considering that the parameters of claim numbers and amounts can be modelled as a function of explanatory variables.c)Proposed calculation of premiums using the variance principle since its takes into account all important information that an insurer knew before the insured join the portfolio and their characteristics for both claim number and amount component.The two component mixture component models developed addresses the bonus hunger problem in that it provides an option to combinations of heavy tailed and light tailed distribution which can give premiums which are tailor made and equitably penalize additionally for claim sizes and lower for small claim amounts.

1.7 Significance of the study

The study of Bonus Malus system play a major role in insurance industry in that the insurance companies will be able to charge premiums based on the risk that the policyholder imposes on the pool.Also a part from promoting careful driving between policyholders,it helps assess individual risks more accurately through priori risk classifications that on a long term basis every policyholder pay a premium which is consistent to their own claim frequency and claim severity.

2 Poisson and Exponential Continuous Mixtures

2.1 Introduction

To obtain the frequency and severity component we shall use the Poisson and exponential mixtures respectively.
A brief discussion of mixtures is given below

2.1.1 Definition

Continuous Mixtures in General

$$f(y) = \int_0^b f(y/\nu)g(\nu)d(\nu$$

where
f(y)=the mixed distribution
$f(y/\nu)$= the conditional probability density
function or probability mass function
$g(\nu)$= continuous mixing or prior distribution
Consider Poisson mixed distribution given by

$P_\nu(n/\nu) = \int_0^\infty \frac{\exp(-\nu)\nu^n}{n!}.g(\nu)d\nu \ ,n = 0,1,2,...$

and type II exponential mixture distribution is given by

$f(y) = \int_0^\infty \frac{1}{\nu}\exp(-\frac{y}{\nu})g(\nu)d\nu \ y \geq 0$ for $\nu \geq 0$

To obtain the mixed distribution,the evaluation of the above integrand explicitly is difficult with the exception of a few mixing distribution,(Albercht,1984).The problem is to obtain other methods of constructing distributions of Poisson mixtures for different probability density functions of $\Lambda = \nu$, and also to identify the ones where explicit evaluation is possible.

The major problem in constructing or obtaining mixture distribution with continuous mixing distribution is the evaluation of the integral above.Only a few integrands can be evaluated explicitly.

2.1.2 The rth Moment

$$
\begin{aligned}
E(x^r) =& \sum_{0}^{\infty} y^r f(x) \\
=& \sum_{a}^{\infty} y^r \int_0^{\infty} f(y/\nu)d(\nu) \\
=& \int_0^{\infty} \{\sum_{0}^{\infty} y^r f(y/\nu)\} g(\nu)d(\nu) \\
=& \int_0^{\infty} E(Y^r/\nu)g(\nu)d(\nu) \\
=& EE(Y^r/\nu)
\end{aligned}
$$

or

$$
\begin{aligned}
E(Y^r) =& \int_0^{\infty} y^r f(y)d(y) \\
=& \int_0^{\infty} \{y^r \int_0^{\infty} f(y/\nu)g(\nu)d(\nu)\}d(y) \\
=& \int_0^{\infty} [\int_0^{\infty} y^r f(y/\nu)d(y)]\, g(\nu)d(\nu) \\
=& \int_0^{\infty} E(y^r/\Lambda)g(\nu)d(\nu) \\
E(Y^r) =& EE(y^r/\Lambda = \nu)
\end{aligned}
$$

2.1.3 Posterior Distribution

Consider two variables $Y = y$ and $\Lambda = \nu$

then,

$$
f(y,\nu) = f(y/\nu)g(\nu)
$$

therefore,

$$\int_0^{\infty} f(y,\nu)d(\nu) = \int_0^{\infty} f(y/\nu)g(\nu)d(\nu)$$
$$1 = \frac{\int_0^{\infty} f(y/\nu)g(\nu)d(\nu)}{\int_0^{\infty} f(y/\nu)g(\nu)d(\nu)}$$

Therefore

$$\frac{f(y/\nu)g(\nu)}{\int_0^{\infty} f(y/\nu)g(\nu)d(\nu)}$$

is a pdf (for $\nu > 0, y > 0$)

Let us denote it by $g(\nu/y)$

therefore,

$$g(\nu/y) = \frac{f(y/\nu)g(\nu)}{\int_0^{\infty} f(y/\nu)g(\nu)d(\nu)} \tag{2.1}$$

$$g(\nu/y) = \frac{f(y/\nu)g(\nu)}{f(y)},$$

for $\nu > 0; y > 0$ is a posterior pdf with posterior mean

$$E(\nu/y) = \frac{\int_0^{\infty} \nu f(y/\nu)g(\nu)d(\nu)}{\int_0^{\infty} f(y/\nu)g(\nu)d(\nu)}$$
$$E(\nu/y) = \frac{E[\nu f(y/\nu)]}{E[f(y/\nu)]}$$

The objectives of this section are to:-

(i) Construct poisson and exponential mixtures for different cases of $g(\nu)$.
(ii) Obtain expectation of y and Variance of y.
(iii) Determine posterior distribution of $f(\nu/y)$ and posterior mean $E(\nu/y)$.

2.2 Poisson Mixtures

Mixed Poisson distributions or Poisson Mixtures were developed in order to model data where the fit of the Poisson distribution was not adequate.Such situation occur often, among other fields in Insurance where an analysis of a heterogeneous portfolio has to be made.As already noted in a heterogeneous portfolio,the fundamental risk for all policyholders to incur an accident are not the same and this justifies the generation of a model that will represent the different underlying risks.It is often convenient when constructing mathematical models of complex phenomenon to use familiar and simple distributions to build more complex distributions.

Mixed Poisson distributions has been given great attention by Johnson et al (1993),Panjer and Willmot (1992) and Douglas (1980).In order to define the mixed Poisson distributions,we let k,the random variable denoting the claim numbers of each insured over a fixed time period which is equal to one year be expressed in terms of Poisson distribution (ν) which differs for different insured policyholders.Thus the parameter ν is the observed value of the random variable Λ and each policyholder's basic risk to have an accident is characterized by the unique value Λ for each risk.When the portfolio is large as it is in our case it is logical to assume that Λ conforms to a continuous distribution in the interval [o,∞].The probability density function of Λ is called mixing function .In the actuarial context it is often named as the risk function and will be denoted as $g(\nu)$.

Thus the distribution of the number of each insured over a year,will have a probability of the form,

$$P_\nu(N=n) = \int_0^\infty \frac{\exp(-\nu)\nu^n}{n!}.g(\nu)d\nu,$$

for $n = 0, 1, 2, \ldots$

2.2.1 The rth Moment

Consider the probability mass function

$$p_y(t) = \int_0^\infty \{\frac{\exp(-\nu t)(\nu t)^y}{y!}\} g(\nu) d(\nu)$$

for y=0,1,2,..

$$\begin{aligned} E(Y^r) =& EE(Y^r/\nu) \\ =& E\{\sum_0^\infty y^r \{\frac{\exp(-\nu t)(\nu t)^y}{y!}\} \end{aligned}$$

Therefore

$$\begin{aligned} E(Y) =& E\{\sum_0^\infty y \frac{\exp(-\nu t)(\nu t)^y}{y!}\} \\ =& E\{\exp(-\nu t)(\nu t) \sum_{y-1}^\infty \frac{(\nu t)^{y-1}}{(x-1)!} \\ =& E\{\exp(-\nu t)(\nu t) \exp(\nu t)\} \\ =& E[t\nu] \\ =& tE[\nu] \\ E(Y^2) =& E[x(y-1)+y] \\ =& E[y(y-1)] + E[Y] \end{aligned}$$

Therefore

$$\begin{aligned} E[y(y-1)] =& EE[y(y-1)/\Lambda] \\ =& E\{\sum_0^{\infty} y(y-1)\frac{\exp(-\nu t)(\nu t)}{y!}\} \\ =& E\{\sum_{y=2}^{\infty} \frac{\exp(-\nu t)(\nu t)^y}{(y-2)!}\} \\ =& E\{(\nu t)^2 \sum_{y=2}^{\infty} \frac{(\nu t)^{y-2}}{(y-2)!} \\ =& E\{(\nu t)^2 \exp(-\nu t) \sum_{y=2}^{\infty} \frac{(\nu t)^{y-2}}{(y-2)!} \\ =& E\{(\nu t)^2 \exp(-\nu t) \exp(\nu t)\} \\ =& E[t^2\Lambda^2] \\ E[y(y-1)] =& t^2 E[\Lambda^2] \end{aligned}$$

Therefore

$$E(Y^2) = t^2 E[\Lambda^2] + tE[\Lambda]$$

Therefore variance is;

$$Var(y) = t^2 E(\Lambda^2) + tE(\Lambda) - t^2[E(\Lambda)]^2$$

$$Var(y) = t^2 Var(\Lambda) + tE(\Lambda) \qquad \textbf{(2.2)}$$

2.2.2 Gamma Mixing Distribution

The mixed model in the portfolio is obtained when,

$k/\nu \sim poisson(\nu)$

and

$\nu \sim Gamma(\alpha, \beta)$

therefore,

$$
\begin{aligned}
p(n) &= \int_0^\infty \frac{\exp(-\nu)\nu^n}{n!} g(\nu) d\nu \\
&= \int_0^\infty \frac{\exp(-\nu)\nu^n}{n!} . \frac{\nu^{(\alpha-1)}\beta^\alpha \exp(-\beta\nu)}{\Gamma(\alpha)} d\nu \\
&= \frac{\beta^\alpha}{n!\Gamma(\alpha)} \int_0^\infty \nu^n \nu^{(\alpha-1)} \exp(-\nu)\exp(-\beta\nu) d\nu \\
&= \frac{\beta^\alpha}{n!\Gamma(\alpha)} \int_0^\infty \nu^{n+\alpha-1} \exp\{-(1+\beta)\nu\} d\nu \\
&= \frac{\beta^\alpha}{k!\Gamma(\alpha)} . \frac{\Gamma(n+\alpha)}{(1+\beta)^{n+\alpha}} \\
&= \frac{\Gamma(n+\alpha}{n!\Gamma(\alpha)} . \frac{\beta^\alpha}{(1+\beta)^{n+\alpha}} \\
P(n) &= \binom{n+\alpha-1}{n} (\frac{\beta}{1+\beta})^\alpha (\frac{1}{1+\beta})^n
\end{aligned}
$$

$$
P(n) = \binom{n+\alpha-1}{n} (\frac{\beta}{1+\beta})^\alpha (\frac{1}{1+\beta})^n \tag{2.3}
$$

Which is the probability density function of the Negative binomial distribution

Mean

$$
\begin{aligned}
E(N) &= \sum_{n=0} n \binom{n+\alpha-1}{n} (\frac{\beta}{1+\beta})^{\alpha} (\frac{1}{1+\beta})^{n} \\
&= \sum_{n=0} n \frac{(n+\alpha-1)!}{n!(\alpha-1)!} (\frac{\beta}{1+\beta})^{\alpha} (\frac{1}{1+\beta})^{n} \\
&= \sum_{n=0} \frac{(n+\alpha-1)!}{(n-1)!(\alpha-1)!} (\frac{\beta}{1+\beta})^{\alpha} (\frac{1}{1+\beta})^{n} \\
&= \frac{1+\beta}{\beta} * \frac{1}{1+\beta} * \alpha \sum_{n=0} \binom{n+\alpha-1}{n-1} (\frac{\beta}{1+\beta})^{\alpha-1} (\frac{1}{1+\beta})^{n-1} \\
E(N) &= \frac{\alpha}{\beta}
\end{aligned}
$$

$$
E(N) = \frac{\alpha}{\beta} \tag{2.4}
$$

Variance

$$
\begin{aligned}
Var(N) =& E[n(n-1)+n] - [E(n)]^2 \\
=& E[n(n-1)] + E(n) - E[(n)]^2
\end{aligned}
$$

$$
\begin{aligned}
E[n(n-1)] &= \sum_{n=0} n(n-1)\binom{n+\alpha-1}{n}(\frac{\beta}{1+\beta})^{\alpha}(\frac{1}{1+\beta})^{n} \\
E[n(n-1)] &= \sum_{n=0} n(n-1)\frac{(n+\alpha-1)!}{n!(\alpha-1)!}(\frac{\beta}{1+\beta})^{\alpha}(\frac{1}{1+\beta})^{n} \\
&= \sum_{n=0} \frac{(n+\alpha-1)!}{(n-2)!(\alpha-1)!}(\frac{\beta}{1+\beta})^{\alpha}(\frac{1}{1+\beta})^{n} \\
&= \sum_{n=0} \alpha(\alpha+1)\frac{(n+\alpha-1)!}{(n-2)!(\alpha-1)!}(\frac{\beta}{1+\beta})^{\alpha}(\frac{1}{1+\alpha})^{n} \\
&= \frac{\alpha(\alpha+1)}{(\beta)^2}\sum_{n=0}\binom{n+\alpha-1}{n-2}(\frac{\beta}{1+\beta})^{\alpha-2}(\frac{1}{1+\alpha})^{n-2} \\
E[n(n-1)] &= \frac{\alpha(\alpha+1)}{(\beta)^2} \\
Var(n) &= \frac{\alpha(\alpha+1)}{(\beta)^2}+\frac{\alpha}{\beta}-(\frac{\alpha}{\beta})^2 \\
&= \frac{\alpha(1+\beta)}{(\beta)^2} \\
Var(n) &= \frac{\alpha}{\beta}(1+\frac{1}{\beta})
\end{aligned}
$$

The Posterior Distribution

from equation 2.1, we note that the formula for posterior distribution is given by

$$
g(\nu/y) = \frac{f(y/\nu)g(\nu)}{\int_0^{\infty} f(y/\nu)g(\nu)d(\nu)}
$$

and

$$
g(\nu) = \frac{\nu^{(\alpha-1)}\beta^{(\alpha)}\exp(-\beta\nu)}{\Gamma(\alpha)},
$$

for $\nu > 0, \alpha > 0, \beta > 0$.

Therefore

$$g(\nu/n) = \frac{\frac{\exp(-\nu t)\nu^n}{n!} * \frac{\nu^{(\alpha-1)}\beta^\alpha \exp(-\beta\nu)}{\Gamma(\alpha)}}{\int_0^\infty \frac{\exp(-\nu t)\nu^n}{n!} * \frac{\nu^{(\alpha-1)}\beta^\alpha \exp(-\beta\nu}{\Gamma(\alpha)} d\nu}$$
$$g(\nu/n) = \frac{\exp(-\nu t)\nu^n \nu^{(\nu-1)}\beta^\alpha \exp(-\beta\nu)}{\int_0^\infty \exp(-\nu t)\nu^n \nu^{(\alpha-1)}\beta^\alpha \exp(-\beta\nu)d\nu}$$
$$g(\nu/n) = \frac{\nu^{n+\alpha-1}\exp(-(t+\beta)\nu)}{\int_0^\infty \nu^{n+\alpha-1}\exp(-(t+\beta)\nu)d\nu}$$
$$g(\nu/n) = \frac{\nu^{n+\alpha-1}\exp(-(t+\beta)\nu)}{\frac{\Gamma(k+\alpha)}{(t+\beta)^{n+\alpha}}}$$
$$g(\nu/n) = \frac{(t+\beta)^{n+\alpha}\nu^{n+\alpha-1}\exp(-(t+\beta)\nu)}{\Gamma(\alpha+n)}$$

$$g(\nu/n) = \frac{(t+\beta)^{\alpha+n}\nu^{n+\alpha-1}\exp(-(t+\beta)\nu)}{\Gamma(\alpha+n)} \tag{2.5}$$

Posterior Mean

$$E(\nu/n) = \int_0^\infty \nu \frac{(t+\beta)^{n+\alpha}\exp(-\nu(t+\beta)\nu^{n+\alpha-1}}{\Gamma(\alpha+n)} d\nu$$
$$E(\nu/n) = \int_0^\infty \frac{(t+\beta)^{n+\alpha}\exp(-\nu(t+\beta)\nu^{n+\alpha}}{\Gamma(\alpha+n)} d\nu$$
$$E(\nu/n) = \frac{(t+\beta)^{n+\alpha}}{\Gamma(\alpha+n)} \int_0^\infty \nu^{n+\alpha}\exp(-\nu(t+\beta)d\nu$$
$$E(\nu/n) = \frac{(t+\beta)^{n+\alpha}}{\Gamma(\alpha+n)} \int_0^\infty \nu^{n+\alpha+1-1}\exp(-\nu(t+\beta)d\nu$$
$$E(\nu/n) = \frac{(t+\beta)^{n+\alpha}}{\Gamma(\alpha+n)} * \frac{\Gamma(n+\alpha+1)}{(t+\beta)^{n+\alpha+1}}$$
$$E(\nu/n) = \frac{(t+\beta)^{n+\alpha}}{\Gamma(n+\alpha)} * \frac{(\alpha+n)\Gamma(n+\alpha)}{(t+\tau)^{n+\alpha}(t+\beta}$$

$$E(\nu/n) = \frac{\alpha+n}{t+\beta} \tag{2.6}$$

2.2.3 Lindley Mixing Distribution

Consider the Poisson distribution of the form

$$P_{\nu}(n) = \frac{\exp(-\nu)\nu^n}{n!}$$

$n = 0, 1, 2, 3, ...$

and assume that ν is Lindley distributed with parameter ϑ

$$g(\nu) = \frac{(\vartheta)^2}{(\vartheta+1)}(\nu+1)\exp(-\nu\vartheta)$$

The unconditional distribution of n will be ;

$$\begin{aligned}
p(n) &= \int_0^\infty p(n/\nu)g(\nu)d\nu \\
&= \int_0^\infty \frac{\exp(-\nu)\nu^n}{n!} * \frac{(\vartheta)^2}{(\vartheta+1)}(\nu+1)\exp(-\nu\vartheta)d\nu \\
&= \frac{(\vartheta)^2}{(\vartheta+1)}\int_0^\infty \frac{\exp(-\nu)\nu^n(\nu+1)\exp(-\nu\vartheta)}{n!}d\nu \\
&= \frac{\vartheta^2}{(\vartheta+1)}[\int_0^\infty \frac{\nu^{n+1}\exp(-\nu(1+\vartheta))}{n!}d\nu + \int_0^\infty \frac{\nu^n\exp(-\nu(1+\vartheta))}{n!}d\nu]
\end{aligned}$$

let $Q=\frac{\vartheta^2}{(\vartheta+1)}$,$R=\frac{n+1}{(1+\vartheta)^{n+2}}$,$S=\frac{1}{(1+\vartheta)^{n+1}}$

$$=Q[R\int_0^\infty \frac{\nu^{n+1}\exp(-\nu(1+\vartheta))(1+\vartheta)^{n+2}}{\Gamma(n+2}d\nu+S\int_0^\infty \frac{\nu^{n}\exp(-\nu(1+\vartheta))(1+\vartheta)^{n+1}}{\Gamma(n+1)}d\nu]$$

$$=\frac{\vartheta^2}{(\vartheta+1)}[\frac{n+1}{(\vartheta+1)^{n+2}}+\frac{1}{(\vartheta+1)^{n+1}}]$$

$$=\frac{\vartheta^2}{(\vartheta+1)}[\frac{(n+1)+(\vartheta+1)}{(\vartheta+1)^{n+2}}]$$

$$p(n)=\frac{\vartheta^2(n+2+\vartheta)}{(\vartheta+1)^{n+3}}$$

$$p(n)=\frac{\vartheta^2(n+2+\vartheta)}{(\vartheta+1)^{n+3}} \tag{2.7}$$

Mean

$$E(N)=\int_0^\infty [\sum_{k=0}^{\infty} n\frac{\exp(-\nu)\nu^n}{n!}]*\frac{\vartheta^2}{(\vartheta+1)}*(1+\nu)\exp(-\vartheta\nu)d\nu$$

$$=\int_0^\infty [\sum_{n=0}^{\infty} \frac{\exp(-\nu)\nu^n}{n!}]*\frac{\vartheta^2}{(\vartheta+1)}(1+\nu)\exp(-\vartheta\nu)d\nu$$

$$=\frac{\vartheta^2}{(\vartheta+1)}\int_0^\infty [\sum_{n=1}^{\infty} \frac{\exp(-\nu)\nu^{n-1}}{(n-1)!}](1+\nu)\exp(-\vartheta\nu)d\nu$$

$$=\frac{\vartheta^2}{(\vartheta+1)}\int_0^\infty \nu(1+\nu)\exp(-\vartheta\nu)d\nu$$

$$=\frac{\vartheta^2}{(\vartheta+1)}[\int_0^\infty \nu\exp(-\vartheta\nu)d\nu+\int_0^\infty \nu^2\exp(-\vartheta\nu d\nu]$$

Using integration by parts we have;

$$E(N) = \frac{\vartheta^2}{1+\vartheta}(\frac{2}{\vartheta^3} + \frac{1}{\vartheta^2})$$
$$E(N) = \frac{2+\vartheta}{\vartheta(1+\vartheta)}$$

Variance

$$Var(N) = E[(n(n-1)] + E(n) - [E(N)]^2$$
$$E[n(n-1)] = \int_0^\infty [\sum_{n=0}^\infty n(n-1)\frac{\exp(-\nu)\nu^n}{n!}](\frac{\vartheta^2}{1+\vartheta})(1+\nu)\exp(-\vartheta\nu)d\nu$$
$$= \int_0^\infty [\sum_{n=2}^\infty \frac{\exp(-\nu)\nu^n}{(n-2)!}](\frac{\vartheta^2}{(1+\nu)}(1+\nu)\exp(-\vartheta\nu)d\nu$$
$$= \int_0^\infty [\nu^2 \sum_{k=2}^\infty \frac{\exp(-\nu)\nu^{n-2}}{(n-2)!}](\frac{\vartheta^2}{(1+\nu)}(1+\nu)\exp(-\vartheta\nu)d\nu$$
$$= \frac{\vartheta^2}{1+\vartheta}[\int_0^\infty \nu^2 \exp(-\vartheta\nu)d\nu + \int_0^\infty \nu^3 \exp(-\vartheta\nu)d\nu]$$

Using integration by parts we have;

$$= \frac{\vartheta^2}{1+\vartheta}[\frac{2}{\vartheta^3} + \frac{6}{\vartheta^4}]$$
$$E[n(n-1)] = \frac{2\vartheta+6}{(1+\vartheta)\vartheta}$$
$$Var(N) = \frac{2\vartheta+6}{(1+\vartheta)\vartheta} + \frac{2+\vartheta}{\vartheta(1+\vartheta)} + (\frac{2+\vartheta}{\vartheta(1+\vartheta)}$$
$$Var(N) = \frac{\vartheta^3+4\vartheta^2+8\vartheta+2}{(1+\vartheta)^2)\vartheta^2}$$

The Posterior Distribution

The posterior structure function formula is obtained by formula given by;

$$g(\nu/n) = \frac{p(n/\nu)g(\nu}{\int_0^\infty p(n/\nu)g(\nu)d\nu}$$

Therefore;

$$g(\nu/n) = \frac{\frac{\exp(-\nu t)\nu^n}{n!} * (\frac{\vartheta^2}{1+\vartheta})(1+\vartheta)\exp(-\vartheta\nu)}{\int_0^\infty \frac{\exp(-\nu t)\nu^{\sum_{i=1}^t n_i}}{n!}(\frac{\vartheta^2}{1+\vartheta})(1+\vartheta)\exp(-\vartheta\nu)d\nu}$$

$$g(\nu/n) = \frac{\exp(-\nu t)\nu^n(1+\nu)\exp(-\vartheta\nu)}{\int_0^\infty \exp(-\nu t)\nu^n(1+\nu)\exp(-\vartheta\nu)d\nu}$$

$$g(\nu/n) = \frac{\exp(-\nu t)\nu^n\exp(-\vartheta\nu) + \nu^{n+1}\exp(-\nu t)\exp(-\vartheta\nu)}{\int_0^\infty \exp(-\nu t)\nu^n\exp(-\vartheta\nu)d\nu + \int_0^\infty \nu^{n+1}\exp(-\nu t)\exp(-\vartheta\nu)d\nu}$$

$$g(\nu/n) = \frac{\nu^n\exp(-\nu(t+\vartheta) + \nu^{n_+1)}\exp(-\nu(t+\vartheta)}{\int_0^\infty \nu^n\exp(-\nu(t+\vartheta)d\nu + \int_0^\infty \nu^{n+1}\exp(-\nu(t+\vartheta)d\nu}$$

$$g(\nu/n) = \frac{\nu^k\exp(-\nu(t+\vartheta) + \nu^{n+1}\exp(-\nu(t+\vartheta)}{\frac{\Gamma(n+1)}{(t+\vartheta)^{n+1}} + \frac{\Gamma(n+2)}{(t+\vartheta)^{n+2}}}$$

$$g(\nu/n) = \frac{\nu^n\exp(-\nu(t+\vartheta) + \nu^{n+1}\exp(-\nu(t+\vartheta))}{\frac{\Gamma(n+2)(t+\vartheta)+\Gamma(n+2)}{(t+\vartheta)^{n+2}}}$$

$$g(\nu/n) = \frac{(t+\vartheta)^{n+2} * \exp(-\nu(t+\vartheta))(\nu^{n+1}+\nu^n)}{\Gamma(n+2)+(t+\vartheta)\Gamma(n+1)}$$

Posterior Mean

$$E(\nu/n) = \int_0^\infty \nu \frac{(t+\vartheta)^{n+2} * \exp(-\nu(t+\vartheta)(\nu^{n+1}+\nu^n)}{\Gamma(n+2)+(t+\vartheta)\Gamma(n+1)} d\nu$$

let $A = \frac{1}{\Gamma(n+2)+(t+\vartheta)\Gamma(n+1)}$

$$E(\nu/n) = A[\int_0^\infty (t+\vartheta)^{k+2} \exp(-\nu(t+\vartheta)\nu^{n+2} d\nu + \int_0^\infty (t+\vartheta)^{n+2} \exp(-\nu(t+\vartheta)\nu^{n+1} d\nu]$$

also let $B = \frac{\Gamma(n+3)}{(t+\vartheta)}$

$$E(\nu/n) = A * B[\int_0^\infty \frac{(t+\vartheta)^{n+3} \exp(-\nu(t+\vartheta)\nu^{n+2}}{\Gamma(n+3)} d\nu + \Gamma(n+2) \int_0^\infty (t+\vartheta)^{n+2} \exp(\nu(t+\vartheta)\nu^{n+1} d\nu]$$

$$E(\nu/n) = A[B+\Gamma(n+2)]$$

substituting for A and B we get;

$$E(\nu/n) = [\frac{1}{\Gamma(n+2)+(t+\vartheta)\Gamma(n+1)}][\frac{\Gamma(n+3)+(t+\vartheta)\Gamma(n+2)}{t+\vartheta}]$$

$$E(\nu/n) = \frac{(n+1)[(n+2)+(t+\vartheta)(n+1)]}{(n+1)(t+\vartheta)[(n+1)+(t+\vartheta)]}$$

$$E(\nu/n) = \frac{(n+1)[(n+2)+(t+\vartheta)]}{(t+\vartheta)[(n+1)+(t+\vartheta]} \quad \textbf{(2.8)}$$

2.3 Exponential Mixtures

Let

$f(y/\nu) = \int_0^\infty \frac{1}{\nu} \exp(-\frac{y}{\nu})$, $x > 0$ and $\nu > 0$

be the conditional type II exponential distribution whose mean is the parameter ν.

Then the probability mass function is given by;

$f_y(t) = \int_0^\infty \frac{1}{\nu t} \exp(-\frac{y}{\nu t}) g(\nu) d(\nu)$

This is the type II exponential mixture ,with $g(\nu)$ as the mixing distribution.

2.3.1 The rth Moment

$E(Y^r) = EE[y^r/\nu]$

$E(Y^r) = E\{\int_0^\infty y^r f(y/\nu t) dy d\nu\}$

$E(Y^r) = E\{\int_0^\infty y^r \frac{1}{\nu t} \exp(-\frac{y}{\nu t}) dy\}$

$E(Y^r) = E\{\frac{1}{\nu t} \int_0^\infty y^r \exp(-\frac{y}{\nu t}) dy\}$

Let $\frac{y}{\nu t} = f \implies y = \nu t f$

$\implies dy = \nu t df$

$E(Y^r) = E\{\frac{1}{\nu t} \int_0^\infty (\nu t u)^r \exp(-u) df\}$

$E(Y^r) = E\{(\nu t)^r \int_0^\infty u^r \exp(-u) df\}$

$E(Y^r) = E\{(\nu t)^r \Gamma(r+1)\}$

$E(Y^r) = E[(\nu t)^r r!]$

put $r = 1$,them we have the first rth moment

$E(Y^r) = E[\Lambda t]$

$$E(Y^r) = tE(\Lambda) \tag{2.9}$$

$E(y^2) = E[y(y-1) + y]$

$E(y^2) = E[y(y-1)] + E(y)$

but

$E[y(y-1)] = EE[y(y-1)/\nu]$

$E[y(y-1)] = E\{y(y-1) \int_0^\infty y^r \frac{1}{\nu t} \exp(-\frac{y}{\nu t}) dy\}$

$E[y(y-1)] = E\{\int_0^\infty \frac{y^2}{\nu t} \exp(-\frac{y}{\nu t})dy - \int_0^\infty \frac{y}{\nu t} \exp(-\frac{y}{\nu t})dy\}$

using integration by parts

$\int_0^\infty \frac{y^2}{\nu t} \exp(-\frac{y}{\nu t})dy$

Let $u = y^2 \implies du = 2dy$ also $dv = \exp(-\frac{y}{\nu t}) \implies v = -\nu t \exp(-\frac{y}{\nu t})$

$\int_0^\infty \frac{y^2}{\nu t} \exp(-\frac{y}{\nu t})dy = \frac{1}{\nu t}\{[-y^2 \nu t]_0^\infty + 2\int_0^\infty y\nu t \exp(-\frac{y}{\nu t})dy\}$

$= 2\int_0^\infty x \exp(-\frac{x}{\nu t})dx$

Let $u = y \implies du = dy$ and $dv = \exp(-\frac{y}{\nu t}) \implies v = -\nu t \exp(-\frac{y}{\nu t})$

$= 2\{[-y - \nu t \exp(-\frac{y}{\nu t})]_0^\infty + \int_0^\infty \nu t \exp(-\frac{y}{\nu t})dx\}$

$= 2\nu t \int_0^\infty \exp(-\frac{y}{\nu t})dy$

$= 2\nu t[-\nu t \exp(-\frac{y}{\nu t})]_0^\infty$

$E\{\int_0^\infty \frac{y^2}{\nu t} \exp(-\frac{y}{\nu t})dy\} = 2(\nu t)^2$

$E\{\int_0^\infty \frac{y^2}{\nu t} \exp(-\frac{y}{\nu t})dy\} = 2t^2 E(\nu^2)$

also

$E\{\int_0^\infty \frac{y}{\nu t} \exp(-\frac{y}{\nu t})dy\}$

Let $u = y \implies du = dy$ and $dv = \exp(-\frac{y}{\nu t}) \implies v = -\nu t \exp(-\frac{y}{\nu t})$

$E\{\int_0^\infty \frac{y}{\nu t} \exp(-\frac{y}{\nu t})dy\} = E(\frac{1}{\nu t})\{[-y\nu t \exp(-\frac{y}{\nu t})]_0^\infty + \int_0^\infty \nu t \exp(-\frac{y}{\nu t})dy\}$

$= E\{\int_0^\infty \exp(-\frac{y}{\nu t})dy\}$

$= E(\nu t)\{[-\exp(-\frac{y}{\nu t})]_0^\infty\}$

$E\{\int_0^\infty \frac{y}{\nu t} \exp(-\frac{y}{\nu t})dy\} = tE(\nu)$

therefore

$E(Y^2) = 2t^2 E(\nu^2) - tE(\nu) + tE(\nu)$

$Var(Y) = 2t^2E(\Lambda^2) - (tE(\Lambda))^2$

$$Var(Y) = t^2[2E(\Lambda^2) - E(\Lambda)^2] \tag{2.10}$$

2.3.2 Inverse-Gamma Mixing Distribution

The inverse gamma is the only distribution whose corresponding exponential mixture is in explicit form.Given

$f(x/y) = \frac{1}{y}\exp(-\frac{x}{y})$ for $x > 0, y > 0$.

and

$$g(y) = \frac{\frac{1}{m}\exp(-\frac{m}{y})}{(\frac{y}{m})^{s+1}\Gamma(s)}$$

$,m > 0, s > 0, y > 0$

The pdf of the mixture is obtained by

$$f(x) = \int_0^\infty \frac{1}{y}\exp(-\frac{x}{y}) * \frac{\frac{1}{m}\exp(-\frac{m}{y})}{(\frac{y}{m})^{s+1}\Gamma(s)}dy$$
$$f(x) = \int_0^\infty \frac{\frac{1}{y}\exp(-\frac{x}{y} * \frac{1}{m}\exp(-\frac{m}{y}}{(\frac{y}{m})^{s+1}\Gamma(s)}dy$$
$$f(x) = \int_0^\infty \frac{1}{y}\exp(-\frac{x}{y}) * \frac{m^s}{\Gamma(s)} * y^{-s-1}\exp(-\frac{m}{y})dy$$
$$f(x) = \frac{m^s}{\Gamma(s)}\int_0^\infty \frac{1}{y^{s+2}}\exp(-\frac{(m+x)}{y})dy$$

Let $z = \frac{1}{y} \implies y = \frac{1}{z}$ and $dy = \frac{-dy}{z^2}$

Therefor

$$
\begin{aligned}
f(x) =& \frac{m^s}{\Gamma(s)} \int_0^\infty z^{s+2} \exp(-(m+x)z) \frac{dz}{z^2} \\
f(x) =& \frac{m^s}{\Gamma(s)} \int_0^\infty z^s \exp(-(m+x)z) dz \\
f(x) =& \frac{m^s}{\Gamma(s)} * \frac{\Gamma(s+1)}{(m+x)^{s+1}} \\
f(x) =& \frac{m^s}{\Gamma(s)} * \frac{s * \Gamma(s)}{(m+x)^{s+1}} \\
f(x) =& \frac{sm^s}{(m+x)^{s+1}}
\end{aligned}
$$

$$
f(x) = sm^s (m+x)^{-(s+1)} \tag{2.11}
$$

for $x > 0, s > 0, m > 0$

This is the Pareto distribution with parameters *s* and *m*.

Mean

$$
\begin{aligned}
E(x) =& \int_0^\infty xsm^s (x+m)^{-s-1} dx \\
=& sm^s \int_0^\infty x(x+m)^{-s-1} dx
\end{aligned}
$$

using integration by parts

$$
\begin{aligned}
E(x) =& sm^s \{ [\frac{-x}{s}(x+m)^{-s}]_0^\infty + \frac{1}{s} \int_0^\infty (x+m)^{-s} dx \} \\
=& sm^s [\frac{1}{s} \frac{(x+s)^{-s+1}}{-(s-1)}]_0^\infty \\
E(x) =& \frac{m}{s-1}
\end{aligned}
$$

$$
E(x) = \frac{m}{s-1} \tag{2.12}
$$

Variance

$$\begin{aligned}Var(x) =& E(x^2) - [E(x)]^2\\ E(x^2) =& \int_0^\infty x^2 s m^s (x+m)^{-s-1} dx\\ =& s m^s \int_0^\infty x^2 (x+m)^{-s-1} dx\end{aligned}$$

using integration by parts

$$\begin{aligned}E(x^2) =& s m^s \{[\frac{x^2(x+m)^{-s}}{-s}]_0^\infty - 2\int_0^\infty \frac{x(x+m)^{-s}}{-s} dx\}\\ =& 2m^s \int_0^\infty \frac{x(x+m)^{-s}}{-s} dx\\ =& 2m^s\{[\frac{-x(x+m)^{-s+1}}{s-1}]_0^\infty + \frac{1}{s-1}\int_0^\infty (x+m)^{-s+1} dx\}\\ =& \frac{2m^s}{s-1}\int_0^\infty (x+m)^{-s+1} dx\\ =& \frac{2m^s}{s-1}[\frac{-(x+m)^{-s+1}}{s-2}]_0^\infty\\ =& \frac{2m^s m^{-s+2}}{(s-1)(s-2)}\\ E(x^2) =& \frac{2m^2}{(s-1)(s-2)}\\ Var(x) =& E(x^2) - [E(x)]^2\\ Var(x) =& \frac{2m^2}{(s-1)(s-2)} - (\frac{m}{s-1})^2\\ Var(x) =& \frac{2m^2}{(s-2)(s-1)^2}\\ Var(x) =& \frac{m^2}{s-1}[\frac{2}{s-2} + \frac{1}{s-1}]; s > 2\end{aligned}$$

Posterior Distribution

$$g(y/x) = \frac{f(x/y)g(y)}{\int_0^\infty f(x/y)g(y)dy}$$

Therefore;

$$f(y/x) = \frac{\frac{1}{y^n}\exp(-\frac{x}{y})g(y)}{\int_0^\infty \frac{1}{y^n}\exp(-\frac{x}{y})g(y)dy}$$

$$f(y/x) = \frac{\frac{1}{y^n}\exp(-\frac{x}{y})\frac{m^s}{\Gamma(s)}\frac{\exp(-\frac{m}{y})}{y^{s+1}}}{\int_0^\infty \frac{1}{y^n}\exp(-\frac{x}{y}\frac{m^s}{\Gamma(s)}\frac{\exp(-\frac{m}{y}}{y^{s+1}}dy}$$

Let $t = \frac{1}{y} \implies y = \frac{1}{t} \implies dy = -\frac{dt}{t^2}$

$$f(y/x) = \frac{\exp(-\frac{m+x}{y})}{y^{n+s+1}\int_0^\infty t^{n+s+1}\exp\{-(m+x)t\}\frac{dt}{t^2}}$$

$$f(y/x) = \frac{\exp(-\frac{m+x}{y})}{y^{n+s+1}\int_0^\infty t^{n+s-1}\exp\{-(m+x)t\}dt}$$

$$f(y/x) = \frac{\exp(-\frac{m+x}{y})}{y^{n+s+1}\frac{\Gamma(n+s)}{(m+x)^{k+s}}}$$

$$f(y/x) = \frac{(m+x)^{k+s}\exp(-\frac{m+x}{y})}{\Gamma(n+s)y^{n+s+1}}$$

$$f(y/x) = \frac{(m+x)^{k+s}}{\Gamma(n+s)}\exp(-\frac{m+x}{y})y^{-(n+s)-1}$$

Which is the Inverse-Gamma $(n+s, m+x)$

Posterior Mean

$$E(y/x) = \int_0^\infty y \frac{(\frac{1}{m+x})\exp(-\frac{(m+x}{y})}{(\frac{y}{m+x})^{n+s+1}\Gamma(n+s)} dy$$

$$E(y/x) = \int_0^\infty \frac{(\frac{1}{m+x})\exp(-\frac{(m+x}{y})}{(\frac{y}{m+x})^{n+s}(\frac{1}{m+x})\Gamma(n+s)} dy$$

$$E(y/x) = \frac{m+x}{\Gamma(n+s)} \int_0^\infty \frac{(\frac{1}{m+x})\exp(-\frac{(m+x}{y})}{(\frac{y}{m+x})^{n+s}} dy$$

$$E(y/x) = \frac{1}{\Gamma(n+s)} \int_0^\infty \frac{\exp(-\frac{(m+x}{y})}{(\frac{y}{m+x})^{n+s}} dy$$

$$E(y/x) = \frac{(m+x)^{k+s}}{\Gamma(n+s)} \int_0^\infty \frac{1}{y^{n+s}} \exp(-\frac{m+x}{y}) dy$$

$$E(y/x) = \frac{(m+x)^{n+s}}{\Gamma(k+s)} \frac{\Gamma(n+s-1)}{(m+x)^{n+s-1}}$$

$$E(y/x) = \frac{m+x}{n+s-1}$$

2.3.3 Lindley Mixing Distribution

The exponential mixture is constructed as follows:

$$g(\nu) = \frac{\vartheta^2}{\vartheta+1}(\nu+1)\exp(-\vartheta\nu), \nu > 0, \vartheta > 0$$

The pdf of the mixture is,

$$
\begin{aligned}
f(x) &= \int_0^\infty \frac{1}{\nu}\exp(-\frac{x}{\nu})\frac{\vartheta^2}{\vartheta+1}(\nu+1)\exp(-\vartheta\nu)d\nu \\
f(x) &= \frac{\vartheta^2}{\vartheta+1}\int_0^\infty (1+\frac{1}{\nu})\exp(-\vartheta\nu-\frac{x}{\nu})d\nu \\
f(x) &= \frac{\vartheta^2}{\vartheta+1}\int_0^\infty (1+\frac{1}{\nu})\exp(-\vartheta(\nu+\frac{x}{\vartheta\nu}))d\nu \\
f(x) &= \frac{\vartheta^2}{\vartheta+1}\{\int_0^\infty \nu^{1-1}\exp(-\vartheta(\nu+\frac{x}{\vartheta\nu}))d\nu+\int_0^\infty \nu^{0-1}\exp(-\vartheta(\nu+\frac{x}{\vartheta\nu}))d\nu\}
\end{aligned}
$$

let,

$\nu = \sqrt{\frac{x}{\vartheta}}p$,therefore $d\nu = \sqrt{\frac{x}{\vartheta}}dp$

$$
\begin{aligned}
f(x) &= \frac{\vartheta^2}{\vartheta+1}\{\sqrt{\frac{x}{\vartheta}}\int_0^\infty p^{1-1}\exp(-\sqrt{\vartheta x}(p+\frac{1}{p}))dp+\int_0^\infty p^{0-1}\exp(-\sqrt{\vartheta x}(p+\frac{1}{p}))dp\} \\
f(x) &= \frac{2\vartheta}{\vartheta+1}\{\sqrt{\frac{x}{\vartheta}}K_1(2\sqrt{\vartheta x})+K_0(2\sqrt{\vartheta x})\}
\end{aligned}
$$

Mean

Using conditional expectation approach,we have

$$\begin{aligned}
E(x^r) =& r!E[\nu^r] \\
E(x^r) =& \frac{\vartheta^2}{\vartheta+1}\{\int_0^\infty \nu^{r+1}\exp(-\vartheta\nu)+\nu^r\exp(-\vartheta\nu)d\nu\} \\
E(x^r) =& \frac{\vartheta^2}{\vartheta+1}\{\frac{\Gamma(r+2)}{\vartheta^{r+2}}+\frac{\Gamma(r+1)}{\vartheta^{r+1}}\} \\
E(x^r) =& \frac{r!}{\vartheta+1}\{\frac{r+1}{\vartheta^r}+\frac{\vartheta}{\vartheta^r}\} \\
E(x^r) =& \frac{r!}{\vartheta^r(\vartheta+1)}(r+\vartheta+1) \\
E(x^r) =& \frac{(r!)^2(r+\vartheta+1)}{\vartheta^r(\vartheta+1)}
\end{aligned}$$

therefore the first moment when $r=1$;

$$E(x) = \frac{\vartheta+2}{\vartheta(\vartheta+1)}$$

Variance

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

Given

$$E(x^r) = \frac{(r!)^2(r+\vartheta+1)}{\vartheta^r(\vartheta+1)}$$

when $r=2$ we have

$$E(x^2) = \frac{4(3+\vartheta)}{\vartheta^2(\vartheta+1)}$$

therefore

$$
\begin{aligned}
Var(x) =& \frac{4(3+\vartheta)}{\vartheta^2(\vartheta+1)} - (\frac{\vartheta+2}{\vartheta(\vartheta+1)})^2 \\
Var(x) =& \frac{9\vartheta+8}{\vartheta^2(\vartheta+1)^2}
\end{aligned}
$$

Posterior Distribution

Given ;

$$
\begin{aligned}
g(\nu/x) =& \frac{f(x/\nu)g(\nu)}{\int_0^\infty f(x/\nu)g(\nu)d\nu} \\
f(\nu/x) =& \frac{\frac{1}{\nu}\exp(-\frac{x}{\nu})\frac{\vartheta^2}{\vartheta+1}(\nu+1)\exp(-\vartheta\nu)}{\int_0^\infty \frac{1}{\nu}\exp(-\frac{x}{\nu})\frac{\vartheta^2}{\vartheta+1}(\nu+1)\exp(-\vartheta\nu)d\nu} \\
f(\nu/x) =& \frac{\frac{1}{\nu}\exp(-\frac{x}{\nu})(\nu+1)\exp(-\vartheta\nu)}{\int_0^\infty \frac{1}{\nu}\exp(-\frac{x}{\nu})(\nu+1)\exp(-\vartheta\nu)d\nu} \\
f(\nu/x) =& \frac{(1+\frac{1}{\nu})\exp(-\vartheta\nu-\frac{x}{\nu})}{\int_0^\infty (1+\frac{1}{\nu})\exp(-\vartheta\nu-\frac{x}{\nu})d\nu} \\
f(\nu/x) =& \frac{(1+\frac{1}{\nu})\exp(-\vartheta\nu-\frac{x}{\nu})}{\int_0^\infty \nu^{1-1}\exp(-\vartheta\nu-\frac{x}{\nu})d\nu + \int_0^\infty \nu^{0-1}\exp(-\vartheta\nu-\frac{x}{\nu})d\nu}
\end{aligned}
$$

3 Modelling Claim numbers and Losses Based on a Posteriori Criteria

3.1 Introduction

In this chapter we are going to design an optimal bonus-malus system that will modify to the policyholders a premium according to claim numbers of they have and same claim size that their claims incur.That is the bigger the size of each claim the bigger is the premium that the policyholder has to pay.

3.2 The frequency Component

3.2.1 Distributions for the Claim number

The following two probability models will be discussed to represent the distribution of the number of claims observed.
These are;-

(i)Poisson-Gamma Distribution
(ii) The Poisson-Lindley Distribution

Following xekalaki(1983),we are going to have a retrospective look in the distribution of claim numbers and their interpretation of their underlying factors.Greenwood and woods (1919) put forward the three hypothesis which have formed the cornerstone for further investigation into the occurrence of accidents.

These hypothesis are:-
(i)The portfolio is homogeneous and all policyholders have the same underlying risk or the same probability to cause an accident.This is to say that the occurrence of a claim from a policyholder constitutes a chance event,it is the result of the pure chance,and it give rise to Poisson distribution.
(ii)The portfolio is diverse and all policyholders have initially the same basic risk to have an accident but this change by each accident sustained. This hypothesis is known as the true contagion hypothesis and leads to what Greenwood and woods called it the biased distribution.
(iii) The portfolio is diverse and all insured individual have constant but different risks to

have an accident.This is known as the apparent contagion and leads to a model known as the accident proneness model.

Under the apparent contagion hypothesis and assuming that the varying underlying risks are distributed according to the Gamma distribution,Greenwood and Woods showed that the distribution of the number of accidents is negative Binomial.A good fit of the negative binomial distribution was then regarded as an indication of heterogeneity in the accident proneness of the portfolio until Irwin(1941),using a result of Mckendrick(1926),derived the negative binomial distribution for a contagion model based on the assumption that the underlying risk or the probability of a policyholder having an accident increases with the number of previous accidents sustained.

Cresswel and Froggat (1936) formulated a forth model that rejects both the concept of accident proneness and the concept of contagion.It is based on the assumption that each policyholder is liable to spell i.e the periods of time during which the policyholder's performance is weak and all of policyholders accidents occur within those spells.The number of accidents within different spells are independent and also its is independent of the number of spells. Kemp(1967) showed that the negative binomial can be given a "spell" interpretation in the context of Poisson distribution generalized by logarithmic distribution.

3.2.2 Posterior Mean Based on Poisson-Gamma Distribution(Negative Binomial Distribution)

Assume that the portfolio is heterogeneous and that all insure individuals have persistent but unequal basic risks to have an accident,that is we assume that each policyholder is having a different accident proneness.
As we have said in the beginning of this chapter,this hypothesis is known as the apparent contagion and leads to a model known in the literature as the "accident proneness" model.We will show below that the negative binomial model is derived under this hypothesis and subsequently use it as a distribution to represent the number of claims and we will assume that it has been derived under the apparent contagion hypothesis.

Consider that the conditional distribution of the number of claims n given the parameter ν is distributed according to Poisson with parameter ν,

$P_{\nu}(n/\nu) = \frac{\exp(-\nu)\nu^n}{n!}$,

where $n = 0, 1, 2, ...$ and $\nu > 0$..

The parameter ν is the observed value of the random variable Λ and it varies from one insured to another .

The accident proneness of each policyholder is characterized by the value of Λ which is distinctive for each risk.The probability density function of Λ is known as the structure function and will be designated by $g(\nu)$.

Let us assume that the structure function conforms to gamma distribution with two parameters,designated by $gamma(\alpha,\beta)$ and is given by the following pdf;

$g(\nu) = \frac{\nu^{\alpha-1}\beta^{\alpha}\exp(-\beta\nu)}{\Gamma(\alpha)}$, for $\nu > 0, \alpha > 0, \beta > 0.$

The two parameter $Gamma(\alpha,\beta)$ has a mean equal to $\frac{\alpha}{\beta}$,variance equal to $\frac{\alpha}{\beta^2}$

The unconditional distribution of the number of claims n denoted as $p(n)$ or p_n in the portfolio is obtained when,

$n/\nu \sim poisson(\nu)$

and

$\nu \sim Gamma(\alpha,\beta)$

from equation 2.3,we have shown that the number of claims n is distributed according to Negative Binomial distribution with probability density function given by;

$P(n) = \binom{n+\alpha-1}{n}(\frac{\beta}{1+\beta})^{\alpha}(\frac{1}{1+\beta})^{n}$

with mean and variance given by $\frac{\alpha}{\beta}$ and $\frac{\alpha}{\beta}(1+\frac{1}{\beta})$ respectively,

$E(n) = \frac{\alpha}{\beta}$

$Var(n) = \frac{\alpha}{\beta}(1+\frac{1}{\beta})$

The variance of the negative binomial exceeds its mean ,a desirable property which is evident in all Poisson mixtures and allow us to deal with data that presents over-dispersion.

Let us consider a policyholder or a class of policyholders that have been under observation for the last t years.

Assume that the total number of claims that a policyholder had in t years denoted as $N = \sum_{i=1}^{t} n_i$,where n_i is the number of claims that an insured had in year $i = 1,2,...t$.

Suppose $g(\nu)$ is the prior distribution for ν,which denotes our subjective belief or the prior information we have about ν.The posterior distribution of the parameter ν for an insured or a group of insured with claim history $n_1,...,n_t$ will be obtained using the Bayes theorem,and when the parameter ν is continuous value,which is the most common situation and it is denoted by $g(\nu/n_1,...,n_t)$.Extending equation 2.5,we get the following;

$$g(\nu/n_1,...,n_t) = \frac{p(n_1,...kn_t/\nu)\mu(\nu)}{\int_0^{\infty} p(n_1,...n_t/\nu) * \mu(\nu)d(\nu}$$

$$p(n_1,...n_t/\nu) = \prod_{i=1}^{t} \frac{\exp(-\nu)\nu^k}{k!}$$

$$p(n_1,...n_t/\nu) = \frac{\exp(-\nu t)\nu^{\sum_{i=1}^{t} n_i}}{\prod_{i=1}^{t} n_i!}$$

but $\sum_{i=1}^{t} n_i = N$

$p(n_1,...n_t/\nu) = \frac{\exp(-\nu t)\nu^N}{\prod_{i=1}^{t} n_i!}$

and

$g(\nu) = \frac{\nu^{(\alpha-1)}\beta^{(\alpha)}\exp(-\beta\nu)}{\Gamma(\alpha)}$, for $\nu > 0, \alpha > 0, \beta > 0.$

Therefore

$$
\begin{aligned}
g(\nu/n_1,...,n_t) =& \frac{\frac{\exp(-\nu t)\nu^N}{\prod_{i=1}^{t} n_i!} * \frac{\nu^{(\alpha-1)}\beta^{\alpha}\exp(-\beta\nu)}{\Gamma(\alpha)}}{\int_0^\infty \frac{\exp(-\nu t)\nu^N}{\prod_{i=1}^{\infty} n_i} * \frac{\nu^{(\alpha-1)}\beta^{\alpha}\exp(-\beta\nu}{\Gamma(\alpha)} d\nu} \\
g(\nu/n_1,...,n_t) =& \frac{\exp(-\nu t)\nu^N \nu^{(\nu-1)}\beta^{\alpha}\exp(-\beta\nu)}{\int_0^\infty \exp(-\nu t)\nu^K \nu^{(\alpha-1)}\beta^{\alpha}\exp(-\beta\nu) d\nu} \\
g(\nu/n_1,...,n_t) =& \frac{\nu^{N+\alpha-1}\exp(-(t+\beta)\nu)}{\int_0^\infty \nu^{N+\alpha-1}\exp(-(t+\beta)\nu) d\nu} \\
g(\nu/n_1,...,n_t) =& \frac{\nu^{N+\alpha-1}\exp(-(t+\beta)\nu)}{\frac{\Gamma(N+\alpha)}{(t+\beta)^{N+\alpha}}} \\
g(\nu/n_1,...,n_t) =& \frac{(t+\beta)^{N+\alpha}\nu^{N+\alpha-1}\exp(-(t+\beta)\nu)}{\Gamma(\alpha+N)}
\end{aligned}
$$

Therefore

$g(\nu/n_1,...,n_t) = \frac{(t+\beta)^{N+\alpha}\nu^{N+\alpha-1}\exp(-(t+\beta)\nu)}{\Gamma(\alpha+N)}$

Which is the pdf of gamma distribution with parameters $(\alpha + N, t + \beta))$

The optimal selection of ν_{t+1} for an individual policyholder with claim history $k_i, ..., k_t$ will enable us obtain the mean of the posterior distribution,that is,

$$
\begin{aligned}
\hat{\nu}_{t+1}(n_1,...,n_t) =& \int_0^\infty \nu \frac{(t+\beta)^{N+\alpha}\exp(-\nu(t+\beta)\nu^{N+\alpha-1}}{\Gamma(\alpha+N)} d\nu \\
\hat{\nu}_{t+1}(n_1,...,n_t) =& \frac{(t+\beta)^{N+\alpha}}{\Gamma(K+\alpha)} * \frac{\alpha + K * \Gamma(N+\alpha)}{(t+\beta)^{N+\alpha}(t+\beta})
\end{aligned}
$$

$$
\hat{\nu}_{t+1}(n_1,...,n_t) = \frac{\alpha+N}{t+\beta} = \bar{\nu}(\frac{\alpha+N}{\alpha+t\bar{\nu}}) \quad \textbf{(3.1)}
$$

where $\bar{\nu} = \frac{\alpha}{\beta}$

It is clearly shown from above that a policyholder that has caused K accidents within a period of t years calls for an update of the gamma parameters from α and β to $\alpha + N$ and $t + \beta$ respectively and this shows that gamma is a conjugate family of Poisson Likelihood.

The net premiums of the Bonus-Malus System obtained in this way can be written in an interesting an useful way.The net premium which is modified by experience and it is equal to posterior mean ;

$$\hat{v}_{t+1}(n_1,...,n_t) = \frac{\alpha+N}{t+\beta} = \bar{v}(\frac{\alpha+N}{\alpha+t\bar{v}})$$

This can be express as a linear combination of the prior premium $\frac{\alpha}{\beta}$, and the observation,$\frac{n}{t}$,that is

posterior mean of the structure function= Z* (the mean of observed data)+ (1-Z)* mean of prior.

or

$\hat{v}_{t+1}(n_1,...,n_t) = Z * \frac{k}{t} + (1-Z)\frac{\alpha}{\beta}$

where Z is known as the credibility factor and is equal to

$$\frac{\alpha+N}{t+\beta} = \bar{v}(\frac{\alpha+N}{\alpha+t\bar{v}})$$

so that the above equation can be true.

That is the net premium of the optimal BMS can be written in terms of Buhlmann credibility model.The Buhlmann credibility model,which denotes that the posterior mean is the weighted average of the observation and the a posterior premium.

The credibility factor Z can be interpreted as the weight that is given to individual experience and it is 0 for $t = 0$, which increases with time and asymptotically tends to 1.

3.2.3 Posterior Mean Based on Poisson-Lindley Distribution

Let us assumed that the number of claims of each insured conforms to a Poisson distribution with parameter v and v is distributed according to a distribution given by Lindley (1958) and (1965).
According to that distribution the parameter v of the Poisson has a distribution function $\mu(v)$ such that

$\mu(v) = \frac{(\vartheta)^2}{\vartheta+1}(v+1)\exp(-v\vartheta)dv$

The underlying risk v is Poisson distributed with p.d.f

$p(n/\nu) = \frac{\exp(-\nu)\nu^n}{n!}$,$n = 0, 1, 2, ...$ and $\nu > 0$

From equation 2.7 we have proved that the unconditional distribution of n is given by ;

$$p(n) = \int_0^\infty p(n/\nu)g(\nu)d\nu$$
$$= \int_0^\infty \frac{\exp(-\nu)\nu^n}{n!} * \frac{(\vartheta)^2}{(\vartheta+1)}(\nu+1)\exp(-\nu\vartheta)d\nu$$
$$p(n) = \frac{\vartheta^2(n+2+\vartheta)}{(\vartheta+1)^{n+3}}$$

According to Sankaran (1970),the Poisson-Lindley distribution has a p.d.f of the following form;

$p_\vartheta(n) = \frac{\vartheta^2(\vartheta+2+n)}{(\vartheta+1)^{n+3}}$,$n = 0, 1, 2, ...$

The mean and variance are given by

$$\frac{2+\vartheta}{\vartheta(1+\vartheta)}$$

and

$$\frac{\vartheta^3+4\vartheta^2+8\vartheta+2}{(1+\vartheta)^2)\vartheta^2}$$

respectively

The posterior structure function for a given policyholder with claim history $n_1, ..., n_t$ is give by;

$$\mu(n_1, ..., n_t) = \frac{p(n_1, ..., n_t/\nu)\mu(\nu}{\int_0^\infty p(n_1, ..., n_t/\nu)\mu(\nu)d\nu}$$
$$p(n_1, ..., n_t/\nu) = \prod_{i=1}^{t} \frac{\exp(-\nu)\nu^n}{n!}$$
$$p(n_1, ..., n_t/\nu) = \frac{\exp(-\nu t)\nu^N}{\prod_{i=1}^{t} n!}$$

Therefore;

$$\mu(\nu/n_1,...,n_t) = \frac{\frac{\exp(-\nu t)\nu^N}{\prod_{i=1}^t n!} * (\frac{\vartheta^2}{1+\vartheta})(1+\vartheta)\exp(-\vartheta\nu)}{\int_0^\infty \frac{\exp(-\nu t)\nu^K}{\prod_{i=1}^t n!}(\frac{\vartheta^2}{1+\vartheta})(1+\vartheta)\exp(-\vartheta\nu)} d\nu$$

$$= \frac{\exp(-\nu t)\nu^N(1+\nu)\exp(-\vartheta\nu)}{\int_0^\infty \exp(-\nu t)\nu^N(1+\nu)\exp(-\vartheta\nu)d\nu}$$

$$= \frac{\exp(-\nu t)\nu^N \exp(-\vartheta\nu) + \nu^{N+1}\exp(-\nu t)\exp(-\vartheta\nu)}{\int_0^\infty \exp(-\nu t)\nu^N \exp(-\vartheta\nu)d\nu + \int_0^\infty \nu^{N+1}\exp(-\nu t)\exp(-\vartheta\nu)d\nu}$$

$$= \frac{\nu^N \exp(-\nu(t+\vartheta) + \nu^{N+1)}\exp(-\nu(t+\vartheta)}{\int_0^\infty \nu^N \exp(-\nu(t+\vartheta)d\nu + \int_0^\infty \nu^{K+1}\exp(-\nu(t+\vartheta)d\nu}$$

$$= \frac{\nu^N \exp(-\nu(t+\vartheta) + \nu^{N+1}\exp(-\nu(t+\vartheta)}{\frac{\Gamma(N+1)}{(t+\vartheta)^{N+1}} + \frac{\Gamma(N+2)}{(t+\vartheta)^{N+2}}}$$

$$= \frac{\nu^N \exp(-\nu(t+\vartheta) + \nu^{N+1}\exp(-\nu(t+\vartheta)}{\frac{\Gamma(N+2)(t+\vartheta)+\Gamma(N+2)}{(t+\vartheta)^{K+2}}}$$

$$\mu(\nu/n_1,...,n_t) = \frac{(t+\vartheta)^{N+2} * \exp(-\nu(t+\vartheta)(\nu^{N+1}+\nu^N}{\Gamma(N+2)+(t+\vartheta)\Gamma(N+1)}$$

The optimal selection of v_{t+1} will give the mean of posterior distribution which is given by;

$$\hat{v}_{t+1}(n_1,...,n_t) = \int_0^\infty v\frac{(t+\vartheta)^{N+2} * \exp(-v(t+\vartheta)(v^{N+1}+v^N}{\Gamma(N+2)+(t+\vartheta)\Gamma(N+1)}dv$$

let $A = \frac{1}{\Gamma(N+2)+(t+\vartheta)\Gamma(N+1)}$

$$\hat{v}_{t+1}(n_1,...,n_t) = A[\int_0^\infty (t+\vartheta)^{N+2}\exp(-v(t+\vartheta)v^{N+2}dv + \int_0^\infty (t+\vartheta)^{N+2}\exp(-v(t+\vartheta)v^{N+1}dv]$$

also let $B = \frac{\Gamma(N+3)}{(t+\vartheta)}$

$$\begin{aligned}\hat{v}_{t+1}(n_1,...,n_t) &= A*B[\int_0^\infty \frac{(t+\vartheta)^{N+3}\exp(-v(t+\vartheta)v^{N+2}}{\Gamma(N+3)}dv + \Gamma(N+2)\int_0^\infty (t+\vartheta)^{N+2}\exp(v(t+\vartheta)v^{N+1}dv] \\ &= A*[B+\Gamma(N+2)]\end{aligned}$$

substituting for A and B ,we get;

$$\begin{aligned}\hat{v}_{t+1}(n_1,...,n_t) &= \frac{1}{\Gamma(N+2)+(t+\vartheta)\Gamma(N+1)}[\frac{\Gamma(N+3)+(t+\vartheta)\Gamma(N+2)}{t+\vartheta}] \\ &= \frac{(N+1)[(N+2)+(t+\vartheta)(N+1)]}{(N+1)(t+\vartheta)[(N+1)+(t+\vartheta)]} \\ \hat{v}_{t+1}(n_1,...,n_t) &= \frac{(N+1)[(N+2)+(t+\vartheta)]}{(t+\vartheta)[(N+1)+(t+\vartheta]}\end{aligned}$$

Posterior mean is

$$\hat{v}_{t+1}(n_1,...,n_t) = \frac{(N+1)[(N+2)+(t+\vartheta)]}{(t+\vartheta)[(N+1)+(t+\vartheta]} \quad (3.2)$$

3.3 Severity Component

3.3.1 Reasons for Taking Into Account the Severity of Each Claim

The obtained when only the number of claims is taken into account has the disadvantage of penalizing claim numbers independently of their severity,that is without taking into account the size of loss that the claim incurs.As a matter of fact,all Bonus-Malus Systems around the world,with the exception of the BMS enforced in Korea,penalize the number of reported accidents in absence of losses that these accidents caused.This means that an insured with a small amount of loss,for example a mere scratch,are forced to pay the equal premium with the policyholders who caused an accident with a big loss,for example a complete destruction of automobile or a serious bodily injury accident.In this sense a Bonus-Malus System that does not take into consideration the size of each claim is not fair and thus a BMS which can separate the policyholders according to the frequency and the severity of their claims should be done. An advantage of the system which considers the claim amounts in its design is that the drivers will report all the accidents caused because they are aware that the amount of loss caused is take in to consideration and that a driver who caused an accident with a small loss will not be punished the same way with someone who caused an accident with a higher amount of loss.In that way the insured will not push for bonuses and the underestimate of the true frequency will be smaller.

Besides the drivers who have claims with big losses are usually doing serious mistakes,such as reckless driving ,egoistic driving,driving with high speed,illegal overtaking,driving under the influence of alcohol,breaking fundamental driving rules such as not observing traffic lights,the right of way and others,in contrast to the accidents which induce small losses and usually are incurred because of moment of inattention.That is the severity of each claim must be penalized not only as an important factor for economic health of insurance but also because because of the good drivers do not have claims and when they have,these claims are with small loses in contrast with the bad drivers which usually have claims with large loss.

3.3.2 Posterior Mean Based on Pareto Distribution

Suppose that a variate x being the claim size and conditional on y is an exponential distribution with mea y and has the form;

$$f(x/y) = \frac{1}{y}\exp(-\frac{x}{y}); x > 0, y > 0.$$

The policyholders in any given portfolio do not have equal expected amount of claim y making the expected amount of claim a random variable and therefore necessitate that it be put in a in form of a distribution.

Further,suppose y has an inverse gamma distribution with parameters s and m given by;

$$g(y) = \frac{\frac{1}{m}\exp(-\frac{m}{y})}{(\frac{y}{m})^{s+1}\Gamma(s)}$$

The mixture has been proved in 2.11 which is a Pareto distribution and is given by;

$$P(X = x) = \int_0^\infty f(x/y) * g(y)dy$$
$$P(X = x) = \int_0^\infty \frac{\frac{1}{y}\exp(-\frac{x}{y}) * \frac{1}{m}\exp(-\frac{m}{y})}{(\frac{y}{m})^{s+1}\Gamma(s)}dy$$
$$P(X = x) = sm^s(m+x)^{-(s+1)}; x > 0, s > 0, m > 0$$

The Pareto distribution is preferred to exponential distribution in modelling claims severity because the relatively tamed exponential distribution get transformed in to heavy tailed Pareto distribution which make it easy to fit into claim data.The heterogeneity that characterize claim size data from different policyholders is incorporated into the model by assuming that the expected claim amount y is distributed according to inverse gamma.

The details we have for each insured on claim size for the time he was in the portfolio would enable us obtain the posterior distribution of the expected claim amount y that would lead us to obtain a tariff structure that incorporate the expected amount of each claim.

Consider a policyholder who has been under observation for a period of t years.

Let x_n denote the size of claim for the k^{th} claim,where $n = 1, 2, 3, ..., N$ and $\sum_{n=1}^{N} x_n$ is the total claim amount for a policyholder who has been under observation for a period of t years.

Also let us denote n_i the number of claim the policyholder had in year i and $N = \sum_{i=1}^{t} N_i$ denotes the total number of claims he has in the portfolio for t years.

Using the Bayes theorem, the posterior distribution of the claim amount y given the claim amount history of the policyholder $x_i, ..., x_n$ can be obtained as below;

$$g(y/x_1,...,x_n) = \frac{f(x_1,...,x_n/y)g(y)}{\int_0^\infty f(x_1,...,x_n/y)g(y)dy}$$
$$= \frac{\left[\prod_{n=1}^N f(x_n/y)\right]g(y)}{\int_0^\infty \left[\prod_{n=1}^N f(x_n/y)\right]g(y)dy}$$

but

$$\prod_{n=1}^N f(x_n/y) = \prod_{n=1}^N \frac{1}{y}\exp(-\frac{x_n}{y})$$
$$\prod_{n=1}^N f(x_n/y) = \frac{1}{y^N}\exp(-\frac{\sum_{n=1}^N x_n}{y})$$

therefore;

$$f(y/x_1,...,x_n) = \frac{\frac{1}{y^N}\exp(-\frac{\sum_{n=1}^N x_n}{y})g(y)}{\int_0^\infty \frac{1}{y^N}\exp(-\frac{\sum_{n=1}^N x_n}{y})g(y)dy}$$
$$= \frac{\frac{1}{y^N}\exp(-\frac{\sum_{n=1}^N x_n}{y})\frac{m^s}{\Gamma(s)}\frac{\exp(-\frac{m}{y})}{y^{s+1}}}{\int_0^\infty \frac{1}{y^N}\exp(-\frac{\sum_{n=1}^N x_n}{y}\frac{m^s}{\Gamma(s)}\frac{\exp(-\frac{m}{y}}{y^{s+1}}dy}$$

Let $t = \frac{1}{y} \implies y = \frac{1}{t} \implies dy = -\frac{dt}{t^2}$

$$
\begin{aligned}
f(y/x_1,...,x_n) &= \frac{\exp-\frac{m+\sum_{n=1}^{N}x_n}{y}}{y^{N+s+1}\int_0^\infty t^{N+s+1}\exp\{-(m+\sum_{n=1}^{N}x_n)t\}\frac{dt}{t^2}} \\
&= \frac{\exp-\frac{m+\sum_{n=1}^{N}x_n}{y}}{y^{N+s+1}\int_0^\infty t^{N+s-1}\exp\{-(m+\sum_{n=1}^{N}x_n)t\}dt} \\
&= \frac{\exp-\frac{m+\sum_{n=1}^{N}x_n}{y}}{y^{N+s+1}\frac{\Gamma(N+s)}{(m+\sum_{n=1}^{N}x_n)^{N+s}}} \\
&= \frac{(m+\sum_{n=1}^{N}x_n)^{N+s}\exp(-\frac{m+\sum_{n=1}^{N}x_n}{y})}{\Gamma(N+s)y^{N+s+1}} \\
f(y/x_1,...,x_n) &= \frac{(m+\sum_{n=1}^{N}x_n)^{N+s}}{\Gamma(N+s)}\exp(-\frac{(m+\sum_{n=1}^{N}x_n}{y})y^{-(N+s)-1}
\end{aligned}
$$

This is Inverse-Gamma $(N+s, m+\sum_{n=1}^{N}x_n)$

This pdf can be re-written as

$$
\begin{aligned}
f(y/x_1,...,x_n) &= \frac{(m+\sum_{n=1}^{N}x_n)^{N+s}\exp(-\frac{(m+\sum_{n=1}^{N}x_n}{y})}{\Gamma(N+s)y^{N+s+1}} \\
&= \frac{(m+\sum_{n=1}^{N}x_n)^{N+s+1}\exp(-\frac{(m+\sum_{n=1}^{N}x_n}{y})}{m+\sum_{n=1}^{N}x_i y^{N+s+1}\Gamma(N+s)} \\
f(y/x_1,...,x_n) &= \frac{(\frac{1}{m+\sum_{i=1}^{N}x_n})\exp(-\frac{(m+\sum_{n=1}^{N}x_n}{y})}{(\frac{y}{m+\sum_{i=1}^{N}x_n})^{N+s+1}\Gamma(N+s)}
\end{aligned}
$$

From the above it is clearly shown that the Inverse Gamma distribution is a conjugate prior with the exponential likelihood.This is because a policyholder that has made K claims within a period of t years with a total claim amount equal to $\sum_{n=1}^{N}x_n$ This implies that an insured with K claims in any given period with total claim amount equal to $\sum_{n=1}^{N}x_n$ demands an update of the parameters Inverse Gamma distribution from s and m to $N+s$ and $m+\sum_{n=1}^{N}x_n$ respectively.

Optimal selection of y_{t+1} for an insured who has claim amounts x_n $,n = 1,2,3,...N)$ in any given period is estimated as ;

$$\hat{y}_{t+1}(x_1,...,x_n) = \int_0^\infty y \frac{(\frac{1}{m+\sum_{n=1}^N x_n})\exp(-\frac{(m+\sum_{n=1}^N x_n}{y})}{(\frac{y}{m+\sum_{i=1}^K x_n})^{N+s+1}\Gamma(N+s)} dy$$

$$\hat{y}_{t+1}(x_1,...,x_n) = \int_0^\infty \frac{(\frac{1}{m+\sum_{i=1}^N x_n})\exp(-\frac{(m+\sum_{n=1}^N x_n}{y})}{(\frac{y}{m+\sum_{n=1}^N x_n})^{K+s}(\frac{1}{m+\sum_{n=1}^N x_n})\Gamma(K+s)} dy$$

$$\hat{y}_{t+1}(x_1,...,x_n) = \frac{m+\sum_{n=1}^N x_n}{\Gamma(N+s)} \int_0^\infty \frac{(\frac{1}{m+\sum_{n=1}^N x_n})\exp(-\frac{(m+\sum_{n=1}^N x_n}{y})}{(\frac{y}{m+\sum_{n=1}^N x_n})^{N+s}} dy$$

$$\hat{y}_{t+1}(x_1,...,x_n) = \frac{1}{\Gamma(N+s)} \int_0^\infty \frac{\exp(-\frac{(m+\sum_{n=1}^N x_n}{y})}{(\frac{y}{m+\sum_{n=1}^N x_k})^{N+s}} dy$$

$$\hat{y}_{t+1}(x_1,...,x_n) = \frac{(m+\sum_{n=1}^N x_n)^{N+s}}{\Gamma(N+s)} \int_0^\infty \frac{1}{y^{N+s}} \exp(-\frac{(m+\sum_{n=1}^N x_n}{y}) dy$$

$$\hat{y}_{t+1}(x_1,...,x_n) = \frac{(m+\sum_{n=1}^N x_n)^{N+s}}{\Gamma(N+s)} \frac{\Gamma(N+s-1)}{(m+\sum_{n=1}^N x_n)^{N+s-1}}$$

$$\hat{y}_{t+1}(x_1,...,x_n) = \frac{m+\sum_{n=1}^N x_n}{N+s-1}$$

$$\hat{y}_{t+1}(n_1,...,x_n) = \frac{m+\sum_{n=1}^N x_n}{N+s-1} \tag{3.3}$$

3.4 Calculation of Premium using the Net Premium Principle

From above,the mean claim frequency if given by equation 3.1 which is;

$$v_{t+1}(n_1,...,n_t) = \frac{\alpha+N}{t+\beta} = \bar{v}(\frac{\alpha+N}{\alpha+t\bar{v}})$$

where

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

also the mean claim size produced a posterior mean equal to;

$$y_{t+1}(x_1,...,x_n) = \frac{m+\sum_{n=1}^{t} x_i}{N+s-1}$$

Thus according to the net premium principle,the premium that any insured person has to pay is;

$$Premium(p) = \frac{\alpha+N}{t+\beta} * \frac{m+\sum_{i=1}^{t} x_i}{s+N-1} \quad \textbf{(3.4)}$$

To be able to calculate premium using 3.4, the following are required;

(i)the years t that an insured is under our observation.
(ii)the observed total claim numbers $N = \sum_{j=1}^{t} n_j$, where n_j the number of accidents that the policyholder has in the t years.
(iii) total claim amount or the aggregate claim amount.
(iv) the maximum likelihood estimates of the parameters of the negative binomial distributionα and β .
(v) the maximum likelihood estimates for parameters s and m .

3.5 Characteristics of the Optimal BMS Based on Claim numbers and Claim sizes

(1)With information gathered in the past,each insured person will pay a premium equivalent to his/her claim history and this shows it is fair.

(2) At any given period,the premiums received form all the insured person is constant meaning that its is balanced finally;

$$P = \frac{\alpha}{\beta}\frac{m}{s-1}$$

Considering that the number of claims and the size of loss are independent component,we have shown that mean of negative binomial distribution given by equation 2.4 ,and of the second is the mean of pareto distribution given by equation 2.12 as shown in the equations below;

$$E_{\nu}[\nu] = E[E(\nu/n_1,...,n_t)] = \frac{\alpha}{\beta}$$

and

$$E_y[Y] = E[E(y/x_1,...,x_n)] = \frac{m}{s-1}$$

respectively.
(3) The premiums rely on the number of accidents which an insured underwent and also it depends on the distribution of these accidents over the years.The policyholder will pay a smaller premium if he has done all his claims in one single year in the beginning of his driving career.
(4) The premium is different not only due to the claim frequency but also due to the claim severity.The policyholder who has claim with a small premium in comparison with an insured who underwent an accident with a big loss.
(5) We have seen that the net premium a policyholder is paying is equal to the posterior mean of the structure function based on his personal claim history after being scaled with the mean claim amount of all the policyholders.This premium is determined from the posterior mean and now it is not scaled but it is determined from the amount of claims for every insured.
(6) The claim amount component is introduced which is very important from the insurer point of view because it is the component that is used to determine the expenses of the insurer from the accident and therefore the premiums to be charged.

4 Modelling of Optimal BMS Based on Individual Characteristics for both claim frequency and Severity

4.1 Introduction

Dionne and Vanasse designed an optimal Bonus-Malus System that combine risk classification and experience rating of the individual policyholder based on his claim frequency and individual characteristics .This framework was further extended by Nicholas Frangos and Spyridon Vrontos by designing a generalized Bonus-Malus system that combines both the priori and the a posteriori information based on individual policyholder characteristics with both frequency and severity incorporated.

The main motivation behind the modelling of a generalized BMS model that take into account the information known to the insurer about the insurer before he/she join the portfolio is that several variables affects the distribution of the claim numbers and the size of loss distribution and that premiums should vary from one policyholder to another.The variables that could be used are the age,sex and the place of residence of the policyholder;the age and the cubic capacity of the car.

The premiums formula from the generalized BMS will be obtained from the following multiplicative formula;

$premium = GBM_F * GBM_S$

where
GBM_F this the generalized BMS obtained when only the number of claims is considered and
GBM_S this the generalized BMS obtained when only the claims size is considered .

4.2 The Generalized Negative Binomial Model

The generalized Bonus-malus system which take into account the number of claims GBM_F is constructed according to Dionne and Vanasse (1989,199) as follows;

Suppose an insured i has been observed for t years and that this insured i made claims totalling to N_i^j.This claims is distributed according to Poisson distributed with parameter v_i^j, where N_i^j are independent.The expected number of claims of individual i for period j is the denoted by v_i^j.Also consider that v_i^j is a function of the vector of h individual characteristics and that it is denoted as $C_i^j = (c_{i,1}^j, ..., c_{i,h}^j)$, which constitute distinct a priori rating variables.

let $v_i^j = \exp(c_i^j \tau^j)$

where

τ^j is the vector of coefficients.

The non-negativity of v_i^j is implied from the exponential function

The probability becomes

$Prob(N_i^j = n/v_i^j) = \frac{\exp(-v_i^j)(v_i^j)^N}{N!}$, $n = 0, 1, 2, ...$

but $v_i^j = \exp(c_i^j \tau^j)$

$Prob(N_i^j = n/v_i^j) = \frac{\exp(c_i^j \tau^j)[(c_i^j \tau^j)]^N}{N!}$

The h individual characteristics is assumed to provide sufficient information for ascertaining mean number of claims.The information we have before an individual join our portfolio may not have all the important information for determining the mean number of claims and therefore we introduce a random variable ε_i into the regression component.

To factor in unobserved significant priori information,we introduce the random variable ε and the expected number of claims can be written as

$v_i^j = \exp(c_i^j \tau^j + \varepsilon) =$

$= \exp(c_i^j \tau^j)\mu_i$

where $\mu_i = \exp(\varepsilon)$, resulting in a random variable ν_i^j

$$\begin{aligned}
Prob(N_i^j = n/\nu_i^j) =& \frac{\exp(-\exp(c_i^j\tau^j+\varepsilon)[\exp(c_i^j\tau^j+\varepsilon)]^n}{n!} \\
Prob(N_i^j = n/\nu_i^j) =& \frac{\exp(-\mu_i\exp(c_i^j\tau^j))[\mu_i\exp(c_i^j\tau^j]^n}{n!} \\
Prob(N_i^j = n/\nu_i^j) =& \frac{\exp(-D\mu_i)(D\mu_i)^n}{n!}
\end{aligned}$$

where $D = \exp(c_i^j\tau^j)$ and $\mu_i = \exp(\varepsilon_i)$

therefore

$Prob = (N_i^j/\mu_i) = \frac{\exp(-D\mu_i)(D\mu_i)^n}{n!}$

and

$Prob(N_i = n) = \int_0^\infty \frac{\exp(-D\mu_i)(D\mu_i)^n}{k!} g(\mu_i)d\mu_i$

Suppose that μ_i is distributed according to a gamma distribution with $E(\mu_i) = \frac{\alpha}{\tau}$ and $Var(\mu_i) = \frac{\alpha}{\tau^2}$

this means that

$g(\mu_i) = \frac{\tau^\alpha}{\Gamma(\alpha)}\exp(-\tau\mu_i)\mu_i^{\alpha-1}$, $\mu_i > 0, \alpha > 0, \tau > 0$

then

$$\begin{aligned}
Prob(N_i = n) =& \int_0^\infty \frac{\exp(-D\mu_i)(D\mu_i)^k}{k!}\frac{\tau^\alpha}{\Gamma(\alpha)}\exp(-\tau\mu_i)\mu_i^{\alpha-1}d\mu_i \\
Prob(N_i = n) =& \frac{\tau^\alpha}{\Gamma(\alpha)}\frac{D^n}{n!}\int_0^\infty \mu_i^{n+\alpha-1}\exp(-(D+\tau)\mu_i)d\mu_i \\
Prob(N_i = n) =& \frac{\tau^\alpha}{\Gamma(\alpha)}\frac{D^n}{n!}\frac{\Gamma(n+\alpha)}{(D+\tau)^{n+\alpha}} \\
Prob(N_i = n) =& \frac{\Gamma(n+\alpha)}{n!\alpha}(\frac{\tau}{D+\tau})^\alpha(\frac{D}{D+\tau})^n \\
Prob(N_i = n) =& \binom{\alpha+n-1}{n}(\frac{\tau}{D+\tau})^\alpha(\frac{D}{D+\tau})^n, n = 0,1,2,...
\end{aligned}$$

Let $\alpha = \tau = a$

Therefore,

$Prob = (N_i^j = n) = \binom{a+n-1}{n}(\frac{\tau}{D+a})^a(\frac{D}{D+a})^n$,$n = 0, 1, 2, ...$

This can be re-written as

$Prob = (N_i^j = n) = \binom{a+n-1}{n}\frac{a^a}{(D+a)^{n+a)}}$

That is

$$Prob = (N_i^j = n) = \frac{\Gamma(n+a)}{n!\Gamma(a)}\frac{a^a}{[a+(1+\frac{D}{a})]^{n+a}}$$
$$Prob = (N_i^j = n) = \frac{\Gamma(n+a)}{n!\Gamma(a)}\frac{a^a}{a^{n+a}}(1+\frac{D}{a})^{-(n+a)}$$
$$Prob = (N_i^j = n) = \frac{\Gamma(n+a)}{n!\Gamma(a)}(\frac{D}{a})^n(1+\frac{D}{a})^{-(n+a)}$$
$$Prob = (N_i^j = n) = \frac{\Gamma(k+a)}{n!\Gamma(a)}(\frac{\exp(c_i^j\tau^j)}{a})^n(1+\frac{\exp(c_i^j\tau^j)}{a})^{-(n+a)}$$

As given by Frangos and Vrontos (2001).This is a Negative Binomial distribution with parameters a and $\exp(c_i^j\tau^j)$

To obtain $E[K_i^j]$ and $Var[K_i^j]$, we should note that

$Prob(K_i^j/\mu_i) = \frac{\exp(-D\mu_{i)}(D\mu_i)}{k!}$, $k = 0, 1, 2, ...$

and

$g(\mu_i) = \frac{\tau^\alpha}{\Gamma(\alpha)}\exp(-\tau\mu_i)\mu_i^{\alpha-1}$,$\mu_i > 0, \alpha > 0, \tau > 0.$

From equation 2.2, we have the following formulae;

$E(x) = tE(\Lambda)$ and $Var(x) = t^2Var(\Lambda) + tE(\Lambda)$

In this case $t = D$,and $\Lambda = \mu_i$,$x = N_i^j$

Therefore $E[N_i^j] = DE[\mu_i] = D\frac{\alpha}{\tau}$

and $Var[N_i^j] = D^2Var(\mu_i) + DE(\mu_i)$

Thus, $E[N_i^j] = D = \exp(c_i^j \tau^j)$,since $\alpha = \tau = a$

$$\begin{aligned} Var[N_i^j] =& D^2 \frac{\alpha}{\tau^2} + D\frac{\alpha}{\tau} \\ Var[N_i^j] =& \frac{D^2}{a} + D \\ Var[N_i^j] =& D(1 + \frac{D}{a}) \\ Var[N_i^j] =& [\exp(c_i^j \tau^j)]\,[1 + \frac{\exp(c_i^j \tau^j}{a}] \end{aligned}$$

$$Var[N_i^j] = [\exp(c_i^j \tau^j)]\,[1 + \frac{\exp(c_i^j \tau^j}{a}] \tag{4.1}$$

Consequently,the insurer needs to obtain,at the renewal of the policy, the best estimate of the expected number of claims ,$\hat{v}_i^{t+1}$, at time $t+1$ for an individual policyholder with a claim history $N_i^1, ..., N_i^t$ and $c_i^1, ..., c_i^{t+1}$ individual known characteristics.we denote $\sum_{j=1}^{t} N_i^j$ be the total number of claims that an individual policyholder i had .The mean claim frequency for an individual policyholder i over the time period t+1 is $\hat{v}_i^{t+1}(N_i^1, ..., N_i^t; c_i^1, ..., c_i^{t+1})$,which is a function of both claim history and individual characteristics.

Using Bayes' theorem the posterior distribution for the policyholder with $N_i^1, ..., N_i^t$ claim history and $c_i^1, ..., c_i^{t+1}$ characteristics can be derived as follows.

$\hat{v}_i^{t+1}(N_i^1, ..., N_i^t; c_i^1, ..., c_i^{t+1}) = \int_0^\infty v_i^{t+1}(N_i^{t+1}, \mu_i) f(v_i^{t+1}/N_i^1, ..., N_i^t; c_i^1, ..., c_i^t) dv_i^{t+1}$

Where by Bayes' rule

$$f(v_i^{t+1}/N_i^1, ..., N_i^t; c_i^1, ..., c_i^t) = \frac{p[(N_i^1, ..., N_i^t)/v_i^{t+1}; c_i^1, ..., c_i^t)]\, f(v_i^{t+1})}{\hat{p}[(n_i^1, ..., n_i^t)/c_i^1, ..., c_i^t]}$$

By definition,

$$\hat{p}[(N_i^1, ..., N_i^t)/c_i^1, ..., c_i^t] = \int_0^\infty p[(N_i^1, ..., N_i^t)/v_i^{t+1}; c_i^1, ..., c_i^t)]\, f(v_i^{t+1}) dv_i^{t+1}$$

Then, $\hat{\nu}_i^{t+1}(N_i^1,...,N_i^t;c_i^1,...,c_i^{t+1})$ is the posterior mean of claim frequency for individual i $[\nu_i^{t+1}(N_i^{t+1},\mu_i)]$,which is a function of both random factor μ_i and individual characteristics.

$$p[(N_i^1,...,N_i^t)/\nu_i^{t+1};c_i^1,...,c_i^t)] = \prod_{j=1}^{t} \frac{\exp(-\exp(c_i^j\tau^j\mu_i)[exp(c_i^j\tau^j\mu_i]^{N_i^j}}{N_i^j!}$$

Let $A = exp(c_i^j\tau^j)$

$$p[(N_i^1,...,N_i^t)/\nu_i^{t+1};c_i^1,...,c_i^t)] = \prod_{j=1}^{t} \frac{\exp(-A\mu_i)[A\mu_i]^{N_i^j}}{n_i^j!}$$

$$p[(N_i^1,...,N_i^t)/\nu_i^{t+1};c_i^1,...,c_i^t)] = \frac{\exp(-\mu_i \sum_{j=1}^t A)[\prod_{j=1}^t (A\mu_i)^{N_i^j}}{\prod_{j=1}^t (N_i^j!)}$$

Therefore

$$f(\mu_i^{t+1}/N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{\frac{\exp(-\mu_i \sum_{j=1}^t A)[\prod_{j=1}^t (A\mu_i)^{N_i^j}}{\prod_{j=1}^t (N_i^j!)} g(\mu_i}{\int_0^\infty [\frac{\exp(-\mu_i \sum_{j=1}^t [\prod_{j=1}^t (A\mu_i)^{N_i^j}}{\prod_{j=1}^t (N_i^j!)}] g(\mu_i d\mu_i}$$

$$f(\mu_i^{t+1}/N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{\exp(-\mu_i \sum_{j=1}^t A)\mu_i^{\sum_{j=1}^t N_i^j} g(\mu_i)}{\int_0^\infty \exp(-\mu_i \sum_{j=1}^t A)\mu_i^{\sum_{j=1}^t N_i^j} g(\mu_i) d\mu_i}$$

$$g(\mu_i) = \frac{\tau^\alpha}{\Gamma(\alpha)} \mu_i^{\alpha-1} \exp(-\tau\mu_i)$$

$$f(\mu_i^{t+1}/N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{\exp(-\mu_i \sum_{j=1}^t A)\mu_i^{\sum_{j=1}^t N_i^j} g(\mu_i) \frac{\tau^\alpha}{\Gamma(\alpha)} \mu_i^{\alpha-1} \exp(-\tau\mu_i)}{\int_0^\infty \exp(-\mu_i \sum_{j=1}^t A)\mu_i^{\sum_{j=1}^t N_i^j} \frac{\tau^\alpha}{\Gamma(\alpha)} \mu_i^{\alpha-1} \exp(-\tau\mu_i) d\mu_i}$$

$$f(\mu_i^{t+1}/N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{\mu_i^{\alpha+\sum_{j=1}^t N_i^j - 1} \exp[-(\tau + \sum_{j=1}^t A)\mu_i]}{\int_0^\infty \mu_i^{\alpha+\sum_{j=1}^t N_i^j - 1} \exp[-(\tau + \sum_{j=1}^t A)\mu_i] d\mu_i}$$

$$f(\mu_i^{t+1}/N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{\mu_i^{\alpha+\sum_{j=1}^t N_i^j - 1} \exp[-(\tau + \sum_{j=1}^t A)\mu_i]}{\frac{\Gamma(\alpha + \sum_{j=1}^t N_i^j)}{(\tau + \sum_{j=1}^t A)^{\alpha + \sum_{j=1}^t N_i^j}}}$$

$$f(\mu_i^{t+1}/N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{(\tau + \sum_{j=1}^t A)^{\alpha + \sum_{j=1}^t N_i^j} \mu_i^{\alpha+\sum_{j=1}^t N_i^j - 1} \exp[-(\tau + \sum_{j=1}^t A)\mu_i]}{\Gamma(\alpha + \sum_{j=1}^t N_i^j)}$$

Let $\alpha = \tau = a$ and $A = \exp(c_i^j \tau^j)$

$$f(\mu_i^{t+1}/N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{(a + \sum_{j=1}^t \exp(c_i^j \tau^j))^{a + \sum_{j=1}^t k_i^j} \mu_i^{a+\sum_{j=1}^t N_i^j - 1} \exp[-(\tau + \sum_{j=1}^t \exp(c_i^j \tau^j))\mu_i]}{\Gamma(a + \sum_{j=1}^t N_i^j)}$$

Which is gamma with updated parameters $[a + \sum_{j=1}^t N_i^j, a + \sum_{j=1}^t \exp(c_i^j \tau^j)]$

The optimal estimate given the observation of $N_i^1,...,N_i^t;c_i^1,...,c_i^t$ is equal to

$$\hat{\mu}_i^{t+1}(N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \int_0^\infty \mu_i \frac{(a+\sum_{j=1}^t \exp(c_i^j\tau^j))^{a+\sum_{j=1}^t N_i^j} \mu_i^{a+\sum_{j=1}^t N_i^j - 1} \exp[-(\tau+\sum_{j=1}^t \exp(c_i^j\tau^j))\mu_i]}{\Gamma(a+\sum_{j=1}^t N_i^j)} d\mu_i$$

$$\hat{\mu}_i^{t+1}(N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{(a+\sum_{j=1}^t \exp(c_i^j\tau^j))^{a+\sum_{j=1}^t N_i^j}}{\Gamma(a+\sum_{j=1}^t N_i^j)} \int_0^\infty \mu_i^{a+\sum_{j=1}^t N_i^j} \exp[-(a+\sum_{j=1}^t \exp(c_i^j\tau^j))\mu_i]\, d\mu_i$$

$$\hat{\mu}_i^{t+1}(N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{(a+\sum_{j=1}^t \exp(c_i^j\tau^j))^{a+\sum_{j=1}^t N_i^j}}{\Gamma(a+\sum_{j=1}^t N_i^j)} \frac{\Gamma(a+\sum_{j=1}^t N_i^j+1)}{(a+\sum_{j=1}^t exp(c_i^j\tau^j)^{a+\sum_{j=1} N_i^j+1}}$$

$$\hat{\mu}_i^{t+1}(N_i^1,...,N_i^t;c_i^1,...,c_i^t) = \frac{a+\sum_{j=1}^t N_i^j}{a+\sum_{j=1}^t \exp(c_i^j\tau^j)} \quad \textbf{(4.2)}$$

4.3 The Generalized Pareto Model

Consider an individual policyholder i with an experience over a period of t years.Let n_i^j denotes the number of claims of policyholder i for period j, N denotes the total number of claims of policyholder i and $X_{i,n}^j$ to denote the size of loss incurred from his claim n for period j .The claims size history is the information we have about the claims incurred and is in the form of a vector $X_{i,1}, X_{i,2}, ..., X_{i,n}$.The total claim amounts for a policyholder i in our portfolio observed over t periods is equal to $\sum_{n=1}^N X_{i,n}$.

Let us assume that $X_{i,n}^j$ follows an exponential distribution with parameter y_i^j.Where y_i^j is the mean claim severity of an individual policyholder in period j.

Since our portfolio is heterogeneous, all expected claim severity is not the same for all the insured individuals and it is therefore fair that each policyholder pays pays a premium proportional to his/her mean claim severity.Consider that the expected claim severity is a function of the vector of h individual's characteristics,denoted as $d_i^j = (d_{i,1}^j, ..., d_{i,h})$, which represent a priori rating variables.Specifically assume that $y_i^j = exp(d_i^j r^j)$ where r is the vector of the coefficients.The non-negativity of y_i^j is implied from the exponential function.

This probability becomes

$Prob(X_{i,n}^{j}) = \frac{1}{y_i^j}\exp(-\frac{x}{y_i^j})$

$Prob(X_{i,n}^{j}) = \frac{1}{\exp(d_i^j r^j)}\exp(-\frac{x}{\exp(d_i^j r^j)})$

The h individual characteristics is assumed to provides sufficient information for ascertaining the mean claim severity.Nonetheless,if the information we have about an insured before he/she join our portfolio is enough ,we have to introduce a random variable z_i into the regression component.

i.e

$$y_i^j = \exp(d_i^j r^j + z_i)$$
$$y_i^j = \exp(d_i^j) w_i$$

Where $w_i = \exp(z_i)$

Therefore

$$Prob(X_{i,n}^{j} = x) = \frac{1}{\exp(d_i^j r^j + z_i)}\exp(-\frac{x}{\exp(d_i^j r^j + z_i})$$
$$Prob(X_{i,n}^{j} = x) = \frac{1}{\exp(d_i^j r^j) w_i}\exp(-\frac{x}{\exp(d_i^j r^j) w_i})$$
$$Prob(X_{i,n}^{j} = w_i) = \frac{1}{p w_i}\exp(-\frac{x}{p w_i})$$

Where $p = \exp(d_i^j r^j)$

therefore

$Prob(X_{i,n}^{j}) = w_i) = \frac{1}{P w_i}\exp(-\frac{x}{P w_i})$

and

$Prob(X_{i,n}^{j}) = x) = \int_0^\infty \frac{1}{P w_i}\exp(-\frac{x}{P w_i}) g(w_i) dw_i$

Suppose that w_i is distributed according to an Inverse gamma distribution with parameters s and $s-1$ and this distribution has mean and variance given by $E(W_i) = 1$, $Var(W_i) = \frac{1}{s-2}$,$s > 2$ respectively.

That is

$g(w_i) = \frac{\frac{1}{s-1}\exp(-\frac{s-1}{w_i})}{(\frac{w_i}{s-1})^{s+1}\Gamma(s)}$

Then

$$
\begin{aligned}
Prob(X_{i,n}^j = x) &= \int_0^\infty \frac{1}{Pw_i}\exp(-\frac{x}{Pw_i})\frac{\frac{1}{s-1}\exp(-\frac{s-1}{w_i})}{(\frac{w_i}{s-1})^{s+1}\Gamma(s)}dw_i \\
Prob(X_{i,n}^j = x) &= \frac{(s-1)^{s+1}}{P\Gamma(s)(s-1)}\int_0^\infty \frac{1}{w_i}\frac{\exp(-\frac{\frac{x}{P}+(s+1)}{w_i}}{w_i^{s+1}}dw_i \\
Prob(X_{i,n}^j = x) &= \frac{(s-1)^{s+1}}{P\Gamma(s)(s-1)}\frac{\Gamma(s-1)}{(\frac{x}{P}+(s-))^{s+1}} \\
Prob(X_{i,n}^j = x) &= \frac{s(s-1)^s}{\frac{1}{P^s}[x+(s-1)P]^{s+1}} \\
Prob(X_{i,n}^j = x) &= sP^s(s-1)[x+(s-1)P]^{-(s+1)}
\end{aligned}
$$

But $P = \exp(d_i^j r^j)$

$$Prob(X_{i,n}^j = x) = s[(s-1)\exp(d_i^j r^j]^s[x+(s-1)\exp(d_i^j r^j]^{-(s+1)}$$

from equation 2.10,;

$E(X) = tE(\Lambda)$ and $Var(X) = t^2[2E(\Lambda^2) - [E(\Lambda)]^2]$

In this case $t = P$ and $\nu = w_i$,$x = X_{i,n}^j$

Therefore;

$$
\begin{aligned}
E[X_{i,n}^j] &= PE[w_i] \\
E[X_{i,n}^j] &= P * 1 \\
E[X_{i,n}^j] &= P
\end{aligned}
$$

but $P = \exp(d_i^j r^j)$

$$E[X_{i,n}^j] = \exp(d_i^j r^j)$$
$$Var[X_{i,n}^j] = P^2[2 * \frac{1}{s-2} - 1]$$

but $P = \exp(d_i^j r^j)$

$$Var[X_{i,n}^j] = (\exp(d_i^j r^j))^2[\frac{2}{s-2} - 1]$$

This can be re-written as;

$$Var[X_{i,n}^j] = \frac{[(s-1)\exp(d_i^j r^j]^2}{s-1}(\frac{2}{s-2} - \frac{1}{s-1})$$

The main objective is to establish generalized an optimal Bonus-Malus system which take into account individual characteristics and past claim history is that the premium paid per individual policyholder is proportional to his claim size.There insurer needs to determine the expected claim severity at period $t+1$ given that the insured has been in our portfolio for t years and and his/her characteristics observed. .

This estimator is denoted as;

$$\hat{y}^{t+1}(X_{i,1}, ..., X_{i,n}; d_i^j, ..., d_i^{t+1})$$

The joint posterior distribution is given by

$$f(y^{t+1}/X_1, ..., X_n; d_i^1, ..., d_i^{t+1}) = \frac{f(x_1, ..., x_n/y)(g(y)}{\int_0^\infty f(x_1, ..., x_n/y)(g(y)d(y)}$$
$$f(y^{t+1}/X_1, ..., X_n; d_i^1, ..., d_i^{t+1}) = \frac{\prod_{i=1}^n f(x_i/y)g(y)}{\int_0^\infty [\prod_{n=1}^n f(x_i/y)]g(y)dy}$$
$$\prod_{n=1}^N f(x_i/y) = \prod_{n=1}^N (\frac{1}{\exp(d_i^j r^j)z_i}\exp(-\frac{x}{\exp(d_i^j r^j)z_i}))$$

Let $m = \exp(d_i^j r^j)$

$$\prod_{i=1}^{N}\left(\frac{1}{\exp(d_i^j r^j) z_i}\exp\left(-\frac{x}{\exp(d_i^j r^j) z_i}\right)\right) = \prod_{i=1}^{N}\left\{\frac{1}{m z_i}\exp\left(-\frac{x}{m z_i}\right)\right\}$$
$$\prod_{i=1}^{N}\left(\frac{1}{\exp(d_i^j r^j) z_i}\exp\left(-\frac{x}{\exp(d_i^j r^j) z_i}\right)\right) = \left(\frac{1}{m z_i}\right)^N \exp\left(-\frac{\sum_{n=1}^{N} X_{i,n}}{m z_i}\right)$$

Suppose

$$g(z_i) = \frac{\frac{1}{s-1}\exp\left(-\frac{s-1}{z_i}\right)}{\left(\frac{z_i}{s-1}\right)^{s+1}\Gamma(s)}$$

Therefore

$$f(y^{t+1}/X_1,\ldots,X_n;d_i^1,\ldots,d_i^{t+1}) = \frac{\left[\left(\frac{1}{m z_i}\right)^N \exp\left(-\frac{\sum_{n=1}^{N} X_{i,n}}{m z_i}\right)\right]\frac{\frac{1}{s-1}\exp\left(-\frac{s-1}{z_i}\right)}{\left(\frac{z_i}{s-1}\right)^{s+1}\Gamma(s)}}{\int_0^\infty \left[\left(\frac{1}{m z_i}\right)^N \exp\left(-\frac{\sum_{n=1}^{N} X_{i,n}}{m z_i}\right)\right]\frac{\frac{1}{s-1}\exp\left(-\frac{s-1}{z_i}\right)}{\left(\frac{z_i}{s-1}\right)^{s+1}\Gamma(s)} dz_i}$$

$$f(y^{t+1}/X_1,\ldots,X_n;d_i^1,\ldots,d_i^{t+1}) = \frac{\frac{\exp\left\{-\frac{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s+1)}{z_i}\right\}}{\left(\frac{z_i}{s-1}\right)^{s+1} z_i^N}}{\int_0^\infty \frac{\exp\left(-\frac{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s+1)}{z_i}\right)}{\left(\frac{z_i}{s-1}\right)^{s+1} z_i^N} dz_i}$$

$$f(y^{t+1}/X_1,\ldots,X_n;d_i^1,\ldots,d_i^{t+1}) = \frac{\frac{1}{z^{N+s+1}}\exp\left\{-\frac{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)}{z_i}\right\}}{\int_0^\infty \frac{1}{z^{N+s+1}}\exp\left\{-\frac{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s+1)}{z_i}\right\} dz_i}$$

$$f(y^{t+1}/X_1,\ldots,X_n;d_i^1,\ldots,d_i^{t+1}) = \frac{\frac{1}{z^{N+s+1}}\exp\left\{-\frac{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)}{z_i}\right\}}{\frac{\Gamma(N+s)}{\left\{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)\right\}^{N+s}}}$$

$$f(y^{t+1}/X_1,\ldots,X_n;d_i^1,\ldots,d_i^{t+1}) = \frac{\left\{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)\right\}^{N+s}\exp\left\{-\frac{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)}{z_i}\right\}}{z_i^{N+s+1}\Gamma(N+s)}$$

This can be re-written as

$$f(y^{t+1}/X_1,...,X_n;d_i^1,...,d_i^{t+1}) = \frac{\{\frac{\sum_{n=1}^{N} X_{i,n}}{m} + (s-1)\}^{N+s+1} \exp\{-\frac{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)}{z_i}\}}{\{\frac{\sum_{n=1}^{N} X_{i,n}}{m} + (s-1)\} z_i^{N+s+1} \Gamma(N+s)}$$

$$f(y^{t+1}/X_1,...,X_n;d_i^1,...,d_i^{t+1}) = \frac{\{\frac{1}{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)}\} \exp\{-\frac{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)}{z_i}\}}{\{\frac{z_i}{\frac{\sum_{n=1}^{N} X_{i,n}}{m}+(s-1)}\}^{N+s+1} \Gamma(N+s)}$$

Which is Inverse Gamma $[N+s, \frac{\sum_{n=1}^{N} X_{i,n}}{m} + (s-1)]$

This is Inverse Gamma with parameters

$[N+s, (s-1)\exp(d_i^j) + \sum_{n=1}^{N} X_{i,n}]$ as shown by Frangos and Vrontos (2001)

To enable us determine the premiums from our general case of pareto,we need to use Baye's theorem to $\hat{y}^{t+1}$ .This optimal estimator which is the posterior mean will be obtained as follows;

$$\hat{y}^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \int_0^\infty y_i^{t+1}(X_i^{t+1},w_i)f(y_i^{t+1}/X_{i,1},...,X_{i,n};d_i^1,...,d_i^{t+1})dy_i^{t+1}$$

$$\hat{z}_i^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \int_0^\infty z_i^{t+1}(X_i^{t+1},w_i)f(y_i^{t+1}/X_{i,1},...,X_{i,n};d_i^1,...,d_i^{t+1})dz_i^{t+1}$$

$$\hat{z}_i^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \int_0^\infty z_i \frac{\{\frac{1}{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)}\}\exp\{-\frac{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)}{z_i}\}}{\{\frac{z_i}{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)}\}^{N+s+1}\Gamma(N+s)} dz_i^{t+1}$$

$$\hat{z}_i^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \frac{\{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)\}^{N+s}}{\Gamma(N+s)} \int_0^\infty \frac{1}{z_i^{N+s}} \exp\{-\frac{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)}{z_i}\}dz_i^{t+1}$$

$$\hat{z}_i^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \frac{\{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)\}^{N+s}}{\Gamma(N+s)} \int_0^\infty \frac{1}{z_i^{N+s-1+1}} \exp\{-\frac{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)}{z_i}\}dz_i^{t+1}$$

$$\hat{z}_i^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \frac{\{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)\}^{N+s}}{\Gamma(N+s)} \frac{\Gamma(N+s-1)}{\{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)\}^{N+s-1}}$$

$$\hat{z}_i^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \frac{\{\frac{\sum_{n=1}^N X_{i,n}}{\exp(d_i^j r^j)}+(s-1)\}}{N+s-1}$$

$$\hat{z}_i^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \frac{\sum_{n=1}^N X_{i,n}+(s-1)\exp(d_i^j r^j)}{s+N-1}$$

This can be re-written as

$$\hat{z}_i^{t+1}(X_{i,1},...,X_{i,t};d_i^1,...,d_i^{t+1}) = \frac{\sum_{n=1}^N X_i+(s-1)\frac{1}{t}\sum_{j=1}^t \exp(d_i^j r^j)}{s+N-1} \quad \textbf{(4.3)}$$

4.4 Premiums Calculation of the Generalized BMS

Both the severity and the frequency components can be used to obtain the premiums of the generalized Bonus-Malus system.

The product of the generalized optimal BMS based on the frequency and of the generalized BMS based on the severity component will give the premiums of the generalized optimal BMS.

This is given by;

Premium=$GBM_F * GBM_S$

$$Premium = [\frac{a+\sum_{j=1}^{t} N_i^j}{a+\sum_{j=1}^{t} \exp(c_i^j \tau^j)}][\frac{\sum_{n=1}^{N} X_{i,n} + (s-1)\frac{1}{t}\sum_{j=1}^{t} \exp(d_i^j r^j)}{s+N-1}] \tag{4.4}$$

4.5 Characteristics of the Generalized BMS

(1) This system takes into consideration the severity,the important prior rating variables for the claim severity , the number of claims and the significant a prior rating variables for the number of claims,for each policyholder.This mean that the system if fair is charging premiums.
(2)The average premiums charged by the insurer each year is;

$$P = \exp(c_i^{t+1}\tau^{t+1})\exp(d_i^{t+1}r^{t+1})$$

This mean that it is financially balanced.The above equation can be shown by stating the following; equation;

$$E[\hat{v}_i^{t+1}(N_i^1,...,N_i^t;c_i^1,...,c_i^t)] = \exp(c_i^{t+1}\tau^{t+1})$$

and that

$$E[\hat{y}_i^{t+1}(N_i^1,...,N_i^t;d_i^1,...,d_i^t)] = \exp(d_i^{t+1}r^{t+1})$$

This is possible because the number of claims and amounts of claims are independent.
(3) The same premium will be paid by in the beginning by all the policyholders with the

same characteristics which is ;

$$P = \exp(c_i^{t+1}\tau^{t+1})\exp(d_i^{t+1}r^{t+1})$$

(4)The premium will always increase proportionately to accidents numbers and the amount of claim that each extra claim incurred and reduce when there no accidents .
(5) This BMS could result in a decrease in Bonus hunger.This because drivers will be penalized separately depending on the amount of loss arising from the accident.
(6) The expenses of the insurer from the accident is determined principally from the claim severity and its introduction into the design of a BMS is crucial as it helps in calculating the correct premium that must be paid.

4.6 Estimation

The net premium principle equation will be used to calculate premiums i.e

$premium = GBM_F * GBM_S$

The net premium formula for the BMS which take into account information about the insured after he/she join the portfolio is given by equation 3.4.

The net premium formula for the generalized BMS based both on a priori , a posteriori classification criteria and on individual characteristics is given by

$$premium = \frac{1}{t}\sum_{j=1}^{t}\exp(c_i^j\tau^j)[\frac{a+\sum_{j=1}^{t}N_i^j}{a+t\exp(c_i^j\tau^j}][\frac{\sum_{i=1}^{N}x_i+(s-1)\frac{1}{t}\sum_j^t\exp(d_i^j r^j)}{s+N-1}]$$

For us to calculate the premium to be paid,the following has to be considered;

(1) The period t to which an insured is in our portfolio.
(2) The estimate of the parameters c and τ for the case of generalized Negative binomial model.
(3)The estimate of the parameters d and r for the case of generalized Pareto model.
(4) the total claim amount $\sum_{n=1}^{N} x_{i,n}$ and the total number of claims ,$\sum_{n=1}^{t} n_i$ where n_i is the number of accidents caused in year $i = 1, ..., t$.

4.6.1 Estimation of Negative Binomial Distribution parameters

Using methods of Moments

To estimate the parameters by the above method, we equate the sample mean $\bar{n}$ and the sample variance s^2 to the corresponding population values.

That is

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

Therefore

$$\begin{aligned}
\alpha &= \bar{n}\beta \\
Variance(k) = s^2 &= \frac{\alpha}{\beta}(1 + \frac{1}{\beta}) \\
s^2 &= \bar{n}(1 + \frac{1}{\beta}) \\
\frac{1}{\beta} &= \frac{s^2 - \bar{n}}{\bar{n}} \\
\hat{\beta} &= \frac{\bar{n}}{s^2 - \bar{k}} \\
and \\
\hat{\alpha} &= \bar{n}\frac{\bar{n}}{s^2 - \bar{n}} \\
\hat{\alpha} &= \frac{\bar{n}^2}{s^2 - \bar{n}}
\end{aligned}$$

provided $s^2 > \bar{n}$

Using the Maximum Likelihood Method

The maximum likelihood estimation method is more sophisticated than the method of moments and superior even for small sets of data as it gives accurate and sensible estimates.

the maximum likelihood function is defined to be

$$L(\alpha,\beta) = \prod_{n=0}^{\infty} [p_n(\alpha,\beta)]^{D_n}$$

$$Where$$

$$p_n = \binom{n+\alpha-1}{n} (\frac{\beta}{1+\beta})^{\alpha} (\frac{1}{1+\beta})^{n}$$

and D_n is the frequency of n accidents.

$$L(\alpha,\beta) = \prod_{n=0}^{\infty} \frac{(n+\alpha-1)!}{n!} (\frac{\beta}{1+\beta})^{\alpha} (\frac{1}{1+\beta})^{n}$$

$$and$$

$$\hat{\beta} = \frac{\alpha}{\bar{n}}$$

$$n-1 = m$$

$$L(\alpha,\beta) = \prod_{n=0}^{\infty} \frac{(\alpha+m)!}{n!} (\frac{\beta}{1+\beta})^{\alpha} (\frac{1}{1+\beta})^{n}$$

$$L(\alpha,\beta) = \sum_{k=1}^{\infty} D_n [\sum_{m=0}^{n-1} \log(\alpha+m)] - \sum_{n=0}^{\infty} N_k \log n! + \sum_{n=0}^{\infty} D_n [-(\alpha+n)\log(1+\beta) + \alpha \log \beta]$$

The maximum likelihood estimates of α and β will be obtained by taking the partial derivatives of log-likelihood with respect to α and β.

The partial derivative of log-likelihood with respect to α is

$$\frac{\partial l(\alpha,\beta)}{\partial \alpha} = \sum_{n=1}^{\infty} D_n[\sum_{m=0}^{n-1} \frac{1}{\alpha+m}] + \sum_{n=0}^{\infty} D_n[\log \beta] - \log(1+\beta) \sum_{n=0}^{\infty} D_n$$
$$\frac{\partial l(\alpha,\beta)}{\partial \alpha} = \sum_{k=1}^{\infty} D_n[\sum_{m=0}^{n-1} \frac{1}{\alpha+m}] + \sum_{n=0}^{\infty} D_n[\log \beta - \log(1+\beta)]$$

and the partial derivatives of log-likelihood with respect to β is

$$\frac{\partial l(\alpha,\beta)}{\partial \beta} = \sum_{n=0}^{\infty} D_n[-\frac{\alpha+n}{1+\beta}] + \sum_{n=0}^{\infty} D_n(\frac{\alpha}{\beta})$$
$$\frac{\partial l(\alpha,\beta)}{\partial \beta} = \sum_{n=0}^{\infty} [\frac{\alpha}{\beta} - \frac{\alpha+n}{1+\beta}]$$

When we equate the partial derivative of log-likelihood with respect to β we get,

$\sum_{n=0}^{\infty}[\frac{\alpha}{\beta} - \frac{\alpha+n}{1+\beta}] = 0$

$\frac{\alpha}{\beta} = \frac{\alpha+n}{1+\beta}$

$\alpha(1+\beta) = \beta(\alpha+n)$

$\alpha + \alpha\beta = \alpha\beta + \beta n$

$\alpha = \beta n$

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

Equating the partial derivative of log-likelihood with respect to α to zero it is;

$$\begin{aligned}
\frac{\partial l(\alpha,\beta)}{\partial\alpha} &= \sum_{n=1}^{\infty} D_n[\sum_{m=0}^{n-1}\frac{1}{\alpha+m}] + \sum_{n=0}^{\infty} D_n[\log\beta - \log(1+\beta)] \\
\frac{\partial l(\alpha,\beta)}{\partial\alpha} &= nlog\beta - nlog(1+\beta) + \sum_{n=1}^{\infty} D_n[\sum_{m=0}^{n-1}\frac{1}{\alpha+m}] \\
\frac{\partial l(\alpha,\beta)}{\partial\alpha} &= n\log\hat{\beta} - n\log(1+\hat{\beta}) + \sum_{m=0}^{n-1}\frac{1}{\hat{\alpha}+m}] \\
n\log(1+\hat{\beta}) &= n\log\hat{\beta} + \sum_{m=0}^{n-1}\frac{1}{\hat{\alpha}+m}]
\end{aligned}$$

The Newton-Raphson approach will be used in order to solve numerically the above equation.

Replacing $\hat{\beta}$ with $\frac{\alpha}{\bar{n}}$ we have that,

$$H(\hat{\alpha}) = d\log\frac{\hat{\alpha}}{\bar{n}} - d\log(1+\frac{\hat{\alpha}}{\bar{n}}) + \sum_{k=1}^{\infty} D_n[\sum_{m=0}^{n-1}\frac{1}{\hat{\alpha}+m}] = 0$$

and

$$H'\hat{\alpha} = \frac{n\bar{n}}{\hat{\alpha}} - \frac{d\bar{n}}{\bar{n}+\hat{\alpha}} + \sum_{n=1}^{\infty} D_n[\sum_{m=0}^{n-1}\frac{1}{(\hat{\alpha}+m)^2}]$$

The calculation of the maximum likelihood estimator for α at the $v-th$ iteration will be

$$\alpha_v = \alpha_{v-1} - \frac{H(\alpha_{v-1}}{H'(\alpha_{v-1})}$$

A good initial value for α_o is the moment estimate for α, which is

$$\hat{\alpha} = \frac{\bar{n}^2}{s^2-\hat{n}}$$

Iteration are repeated until α_v is sufficiently close to α_{v-1}.After we have obtained $\hat{\alpha}$ we can compute β from the following formula;

$$\hat{\beta} = \frac{\hat{\alpha}}{\bar{n}}$$

Generally the fit of the negative binomial is very good and it give an excellent representation of the drivers behavior.This mean that the Poisson distribution fits well for the individual drivers accidents and the Gamma distribution represents well the heterogeneity of the portfolio,that is the different underlying risk of each policyholders to produce a claim.

4.6.2 Estimation of Pareto Distribution Parameters

Using Methods of Moments

$$\begin{aligned}
E[x] &= \frac{m}{s-1} \\
\bar{x} &= \frac{m}{s-1} \\
m &= \bar{x}(s-1) \\
&and \\
Var[x] = S^2 &= \frac{m^2 s}{(s-2)(s-1)^2} \\
S^2 &= \frac{\hat{x}^2 (s-1)^2 s}{(s-2)(s-1)^2} \\
S^2 &= \frac{\hat{x}^2 s}{(s-2)} \\
\hat{s} &= \frac{2S^2}{S^2 - \hat{x}^2} \\
\hat{m} &= \hat{x}\{\frac{2S^2}{S^2 - \hat{x}^2} - 1\} \\
\hat{m} &= \hat{x}\{\frac{S^2 + \hat{x}^2}{S^2 - \hat{x}^2}\}
\end{aligned}$$

$$\hat{m} = \hat{x}\{\frac{S^2 + \hat{x}^2}{S^2 - \hat{x}^2}\} \tag{4.5}$$

The Maximum Likelihood Method

In order to find the maximum likelihood estimators,we obtain the likelihood which is equal to,

$$L(s,m) = \prod_{i=}^{n} sm^s(x_i+m)^{-s-1}$$

The log-likelihood is given by

$$\log l(s,m) = \sum_{i=}^{n} \ln[sm^s(x_i+m)^{-s-}]$$
$$\log l(s,m) = \sum_{i=}^{n} [\ln s + s\ln m - (s+1)\ln(x_i+m)]$$
$$\log l(s,m) = n\ln s + ns\ln m - (s+1)\sum_{i=1}^{n} \ln(x_i+m)$$

Let $\log l(s,m) = l$

$$l = n\ln s + ns\ln m - (s+1)\sum_{i=1}^{n} \ln(x_i+m)$$

The first derivative of the log-likelihood with respect to s is

$$\frac{\partial l}{\partial s} = \frac{n}{s} + n\ln m - \sum_{i=1}^{n} \ln(x_i+m)$$

The first derivative of log-likelihood with respect to m is

$$\frac{\partial l}{\partial m} = \frac{ns}{m} - (s+1)\sum_{i=1}^{n}\frac{1}{(m+x_i)}$$

The second derivative with respect to s is

$$\frac{\partial^2 l}{\partial s^2} = -\frac{n}{s^2}$$

The first derivative of the log-likelihood with respect to m is

$$\frac{\partial l}{\partial m} = \frac{ns}{m} - (s+1)\sum_{i=1}^{n}\frac{1}{(m+x_i)}$$

The second derivative of the log-likelihood with respect to m is

$$\frac{\partial^2 l}{\partial m^2} = -\frac{ns}{m^2} + (s+1)\sum_{i=1}^{n}\frac{1}{(m+x_i)^2}$$

and the second derivative of the log-likelihood with respect to s and then to m which is

$$\frac{\partial^2 l}{\partial m \partial s} = \frac{\partial^2 l}{\partial s \partial m} = \frac{n}{m} - \sum_{i=1}^{n}\frac{1}{(m+x_i)}$$

We use the method of scoring.The vector of parameters $\vartheta = (s,m)$ and $A(\vartheta)$ as the matrix of the first derivative which is defined as

$$R(\vartheta) = \begin{pmatrix} \frac{\partial l}{\partial s} \\ \frac{\partial l}{\partial m} \end{pmatrix} = \begin{pmatrix} \frac{n}{s} + n\ln m - \sum_{i=1}^{n} \ln(m + x_i) \\ \frac{\partial^2 l}{\partial s \partial m} = \frac{n}{m} - \sum_{i=1}^{n} \frac{1}{(m+x_i)} \end{pmatrix}$$

and the Q(ϑ) the matrix of the second derivatives which is defined as

$$Q(\vartheta) = \begin{pmatrix} \frac{\partial^2 l}{\partial s^2} \frac{\partial^2 l}{\partial s \partial m} \\ \frac{\partial^2 l}{\partial m \partial s} \frac{\partial^2 l}{\partial s^2} \end{pmatrix}$$

$$Q(\vartheta) = \begin{pmatrix} (-\frac{n}{s^2})(\frac{n}{m} - \sum_{i=1}^{n} \frac{1}{m+x_i}) \\ (\frac{n}{m} - \sum_{i=1}^{n} \frac{1}{m+x_i})(-\frac{ns}{m^2} + (s+1)\sum_{i=1}^{n} \frac{1}{(m+x_i)^2}) \end{pmatrix}$$

and then the vector of parameter ϑ is estimated using the following equation

$$\vartheta_{i+1} = \vartheta_i - [Q(\vartheta)]^{-1}[A(\vartheta)]$$

5 Application

5.1 Introduction

We consider the data presented by Spyridon D. Vrontos (2001)

Table 1. Claim Frequency Data.

Number of claims per policy(n)	Number of Policies
0	1755724
1	117632
2	14510
3	2228
4	418
5	73
6	23
7	6
8	1

The mean and variance of this data is obtained as;

$Mean = E(N) = 0.08228$

$Variance = S^2 = 0.1019$

we consider the negative binomial model.

using the method of moments

$$\hat{\beta} = \frac{\bar{n}}{s^2 - \bar{n}}$$
$$\hat{\beta} = 4.1937$$
$$\hat{\alpha} = \frac{\bar{n}^2}{s^2 - \hat{n}}$$
$$\hat{\alpha} = 0.34506$$

using maximum likelihood method

we have that

$$\alpha_v = \alpha_{v-1} - \frac{H(\alpha_{v-1})}{H'(\alpha_{v-1})}$$

A good start for the value of α_0 is the moment estimate for α,which is

$$\hat{\alpha} = \frac{\bar{n}^2}{s^2 - \hat{n}}$$

After we compute α we compute β from the formula

$$\beta = \frac{\hat{\alpha}}{\bar{n}}$$

This after a few iteration we find that $\hat{\alpha} = 0.34854$ and $\hat{\beta} = 4.23602$

5.1.1 Optimal BMS using the Net Premium Principle

Here the optimal BMS is obtained using the net premium principle.All the values obtained are divided with the basic premium,the premium paid when $t = n = 0$, i.e the premium paid in the beginning of the driving career of the policyholder in our portfolio and under observation.We consider this division because in this phase we are more interested in the percentage difference that result after the maluses are given.The results presented indicates that a new policyholder pays premium 100 as the initial premium.

The Negative Binomial Distribution

We apply the Negative Binomial parameter estimates in to equation 3.1 and obtain the Optimal BMS in table 2.It is clear that the Optimal BMS proposed is generous to the good drivers and strict for the bad drivers.It is generous because for the basic premium the bonuses are 19.10 percent for one claim free year,32.07 percent for two claim -free year and the policyholder who had 7 claim free years will have a bonus of 62.3 percent of the basic premium.On the other hand,the driver who had one accident will have increase of 213.017 percent of the basic premium to be paid and the penalties are more severe for drivers who cause two or more accidents in one year.

From this system,it is clear that it increases the phenomenon of hunger bonus the is because the policyholder with a small loss arising from his claim will prefer to pay the damage themselves as they will not want to lose on the big bonuses awarded.

Table 2. Optimal BMS based on the Posteriori Criteria,Negative Binomial Distribution Model

Years	Number of Claims					
t	0	1	2	3	4	5
0	100.00	0.00	0.00	0.00	0.00	0.00
1	80.9015	313.0170	545.1325	777.2480	10009.3635	1241.4790
2	67.9283	262.8221	457.7158	652.6096	847.5034	1042.3972
3	58.5408	226.5007	394.4607	562.4206	730.3806	898.3406
4	51.4329	198.9995	346.5661	494.1328	641.6994	789.2660
5	45.8641	177.4535	309.0428	440.6322	572.2221	703.81081
6	41.3835	160.1173	278.8511	397.5849	516.3188	635.0526
7	37.7004	145.8669	253.0373	362.2001	470.3668	578.5333

5.1.2 Optimal BMS based on the a Posteriori frequency and Severity Component

The independence property has to be assumed between claim numbers and claim amounts in order to deliberate on each component separately.We consider the data presented by Walhin and Paris (2000) and added the values on the right so that the data fits the Pareto distribution.

From table 3 the mean and variance claim amount obtained are;
$E(x) = 321,422.22$
$Var(x) = 2.85637E + 11$

Using the maximum likelihood estimate,we find that $\hat{s} = 3.13$ and $\hat{m} = 685,682.79$.

Here we will demonstrate only the instances that the total claim amount of an insured in the portfolio under our observation i.e $\sum_{n=1}^{N} x_n$ will be equivalent to Kes.250,000.00,Kes.750,000.00 and Kes.1,000,000.00.However we can use different values with the the net premium formula .

In table 4, it shows clearly the premiums that will be paid by insured who underwent different number of claims.For example let us consider an insured who underwent one accident in the first year with a claim amount of Kes.250,000 ,the amount of premium that

Table 3. Observed Claim Severity Distribution '000'

6	6	10	11	17	18	20	26	27	34
42	44	47	54	59	60	61	61	61	61
64	64	65	66	67	68	71	71	73	75
76	81	85	87	93	94	101	103	105	109
110	110	113	116	116	129	134	134	141	141
151	154	156	159	167	171	172	173	174	179
181	183	185	187	195	195	203	226	235	240
151	255	273	340	361	429	465	531	646	923
1043	1226	1398	1423	1569	1702	1929	081	2265	2545

he will pay is Kes.76,992.20.If we observe him for the next period and he made a claim size of Kes.500,000.00 for his accident made, then a surcharge will be enforced as he had an accident and he will have to pay Kes.130,917.90.This premium is equivalent to two accidents observed in a period of two years of total claim amount of Kes.750,000.00.In the event that the insured remain in our portfolio for the third year with no accident,the premium to be paid will be lower and he will pay Kes.112,825.30.This amount has reduced since there was no accident caused and this translates to the total premium for two accidents of total claim amount of Kes.750,000.00 for the three periods the insured is in our portfolio.

Table 4. Optimal BMS Based on the a Posteriori Frequency and Severity Component-Total Claims Size of 250000

Years	Number of Claims					
t	0	1	2	3	4	5
0	26,487.30	0	0	0	0	0
1	21,428.70	76,992.20	101,619.10	116,644.80	126,768.20	134,051.90
2	18,044.90	64,645.90	85,323.60	97,939.80	78,000.70	112,555.50
3	15,505.90	55,712.00	73,532.10	84,404.80	91,730.10	97,000.60
4	13,623.20	48,947.50	64,603.00	74,156.50	80,592.40	85,223.00
5	12,148.20	43,647.90	57,609.20	66,127.50	71,866.50	75,995.80
6	10,961.40	39,383.70	51,981.10	59,667.20	64,845.60	68,571.40
7	9,985.80	35,878.60	47,354.80	56,099.60	59,074.40	62,468.60

Table 5. Optimal BMS Based on the a Posteriori Frequency and Severity Component-Total Claim Size of 750,000.00

Year	Number of Claims					
t	0	1	2	3	4	5
0	26,487.30	0	0	0	0	0
1	21,428.70	118,134.50	155,921.20	178,976.20	194,509.20	205,685.10
2	18,044.90	99,190.60	130,917.90	150,275.80	163,318.00	172,701.70
3	15,505.90	85,482.70	112,825.30	129,508.10	140,747.80	148,834.80
4	13,623.20	75,103.60	99,138.70	113,797.70	123,674.00	130,780.00
5	12,148.20	66,980.40	88,393.70	101,476.70	110,283.60	116,620.20
6	10,961.40	60,436.80	79,768.20	91,563.00	99,509.60	105,227.10
7	9,985.80	55,057.90	72,668.90	83,413.90	90,653.30	95,861.90

Table 6. Optimal BMS based on the a Posteriori Frequency and Severity Component-Total Claim Size of 1000000

Year	Number of Claims					
t	0	1	2	3	4	5
0	26,487.30	0	0	0	0	0
1	21,428.70	138,720.40	183,091.80	210,164.40	228,404.10	241,527.50
2	18,044.90	116,475.40	153,731.40	176,462.70	191,777.50	202,796.50
3	15,505.90	100,378.80	132,486.10	152,076.00	165,274.40	174,770.50
4	13,623.20	88,191.00	116,400.00	133,611.20	145,207.10	153,550.30
5	12,148.20	78,642.40	103,797.10	119,144.90	129,485.30	136,925.10
6	10,961.40	70,959.50	93,656.70	107,505.10	116,835.30	123,548.30
7	9,985.80	64,644.10	85,321.30	97,937.20	106,437.00	112,552.60

In table 7 we illustrates how this bonus-malus system is discriminative of the premium paid when claim amount is taken into account.it also show premium to be paid when an insured is in our portfolio for one year and caused accidents which range from 1 upto 5.The total claim amount range from Kes.250,000.00 to Kes.4,000,000.00. For example an insured who underwent an accident with an amount of Kes.250,000.00 , Kes.750,000.00 and Kes.2,000,000.00 should have to pay a premium of Kes.76,992.20, Kes.118,134.50 and Kes. 220,990.40 respectively.

Table 7. Comparison of Premiums for Various Number of Claims and Claim Sizes in the First Year of Observation

	Number of Claims				
Claim Sizes	1	2	3	4	5
250,000	76,992.20	101,619.10	116,644.80	126,768.20	134,051.90
500,000	97,563.40	128,770.10	147,810.50	160,638.70	169,786.80
1,000,000	138,702.60	183,113.00	210,141.90	228,379.70	241,501.70
2,000,000	220,990.10	291,676.40	334,804.70	363,861.80	384,768.20
3,000,000	303,274.70	400,280.60	459,467.50	499,343.80	528,034.60
4,000,000	385,559.20	508,884.80	584,130.30	634,825.80	671,301.00

5.1.3 The designing of an Optimal BMS based both on the a priori and the a posteriori Classification Criteria

The premiums of a generalized optimal Bonus-Malus system is obtained by apply equation 4.4 .

To enable us implement equation 4.4,which has both negative binomial regression parameters and Pareto parameters,we have have to find the estimate of the parameters a and s ,and the vector τ and r of significant priori rating variables.We employ the methods developed by Riggby and Stasinopoulos (2001,2005,2009) to estimate the parameters in the generalized models.

The generalized additive model for lacation,scale and shape(GAMLSS)

The GAMLSS are models which are semi-parametric regression in nature.The parametric distribution assumption shows that they are parametric and using the non-parametric smoothing functions as function of explanatory variables in modelling of parameters give the sense of "semi".These models were introduced as an option of Generalized Linear Models,GLM and Generalized Additive Models ,GAM, which within the framework of univariate regression modelling techniques hold a prominent place.The response variable y relaxed in the exponential family in the GAMLSS model and replaced by a general distribution family .Within this model the systematic part of the model is expanded to allow not only for the mean(or location) but other parameters of y as,linear and/or non-linear,parametric and/or additive non-parametric (smooth)functions of explanatory variables and/or random effects which is the key advantage of GAMLSS.Two different algorithm were used in model fitting of GAMLSS,the fisrt is based on the algorithm that was used for fitting the mean and dispersion additive models of Rigby and Stasinopoulos,whereas the second is based on the Cole and Green algorithm.

The observations y_i for $i = 1, ..., n$ with probability density function $f(y_i/\tau)$ conditional on $\tau_i = (\tau_{1i}, ..., \tau_{i4)}$ are assumed independent by GAMLSS model,the vector of the distribution parameters each of which can be a function of explanatory variables.

6 Conclusion and Recommendations

In this thesis we presented the methodology for the construction of a BMS according to the number of claims and net premium calculation principle.We used the negative binomial distribution to model the number of claims and we found that the optimal BMS obtained was fair as every policyholder should pay a premium which is in line with the estimate of his/her number of claims.It is finally balanced as each period the total of amount of the premiums recived by the insurance company from all the insured individuals remains constant and can be written as a special case of Buhlmann credibility model which denotes the risk premium modified by experience is the weighted average of the prior premium and the observation.

We have also developed the methodology for designing a BMS using the number and amount of claims components.This optimal BMS adjust to the policyholders a premium according to the number of claims they have and the correct claim size they incur.We used the Negative Binomial distribution as the distribution of the number of claims,the Pareto distribution for the loss that each claim incur and net premium calculation principle.The premium that each policyholder is paying is proportional to the number of his/her claims and to the loss that these claims incur and this indicates that the optimal BMS obtained in this ways is fair.It is finally balanced as each year the mean value of all premiums collected from all the policyholders is stable.This optimal BMS is important because policyholders who are going to pay a small premium are those with a small loss than those with big loss and also for the insurance companies as the size of loss is very important factor for their financial security.

Furthermore,we have developed a generalized BMS for frequency severity model using claim frequency and loss distribution in the design of a tariff classification criteria that take into account both the a priori and a posteriori classification criteria.We have constructed the generalized model for both the frequency and severity component.However,we have not applied this generalized model to the because we needed to estimate the regression parameters and their interaction using the Generalized Additive Model for Location,Scale and Shape (GAMLSS) which we have only described and we recommend this challenge as a subject for further research in estimation of regression components.

6.1 Future Research

As stated in the conclusion above,we recommend the application of GAMLSS in the estimation of regression parameters as a subject for further research.This would enable us apply the generalized models constructed to the data and therefore enabling us to see how premiums adjust to the impact of individual characteristics.This can also be extended to other models using other premium calculation principles.

Finally,we would like to recommend the development the development of a method for the evaluation of the efficiency for the current bonus-malus system when the size of loss and the number of claims is put into consideration as a subject for further research.

References

[DGV] Dionne, G. and Vanasse, C :*A generalization of actuarial automobile insurance rating models:the negative binomial distribution with a regression component,Astin Bulletin 19,199-212 (1989).*

[DGV] Dionne, G. and Vanasse, C :*Automobile insurance ratemaking in the presence of asymmetrical information,journal of Applied Economtrics 7, 149-165 (1992)*

[GMT] Goureroux C., Montfort A. and Trognon A. :*Pseudo maximum likelihood methods:theory,Econometrica 52,681-700 (1984a).*

[GMT] Goureroux C., Montfort A. and Trognon A. :*Pseudo maximum likelihood methods:application to poisson models,Econometrica 52,701-720 (1984b).*

[HBG] Hausmann J.A. Hall B.H. and Griliches Z.:*Econometric models for count data with an application to the patents- Research and Development relationship,Econometria,46,1251-1271. (1984)*

[LJF] Lawless,J.F.:*Negative Binomial Distribution and Mixed Poisson Regression,canadian journal of statistics ,15,3,209-255 (1987).*

[P] Pinquet J.:*Allowance for cost of claims in Bonus-Malus Systems,Austin Bulletin,27,33-57 (1997).*

[PJ] Pinquet J.:*Designing Optimal Bonus-Malus Systems from Different Types of Claims.Austin Bulletin,28,205-220 (1998).*

[RAE] Renshaw A.E.:*Modelling the Claims Process in the presence of covariates.Austin Bulletin,24,265-285 (1994).*

[TL] Tremblay,L.:*Using Poisson Inverse Gaussian in Bonus-Malus Systems,Austin Bulletin 22,97-106 (1992).*

[L] Lemaire,J.:*Bonus-malus systems in automobile insurance,Kluwer Academic Publishers,Massachusets (1995).*

[VS] Vrontos S.:*Design of an Optimal Bonus-Malus System in Automobile Insurance.Msc.Thesis,Department of Statistics,Athen University of Economics and Business,ISB:960-7929-21-07. (1998).*

[FNS] Frangos, N. and S. Vrontos:*Design of Optimal bonus-malus systems with frequency and severity component on an individual basis in automobile insurance.Austin Bulletin,31,1,1-22 (2001).*

[TGF] Tzougas]Tzougas,G. and N. Fraongos:*Design of an optimal bonus-malus systems based on the sichel distribution.collective book.modern problems in insurance mathematics,Springer Verlag (2014).*

[TSF] Tzougas,G.,S. Vrontos and N. Frangos:*optimal bonus-malus systems using finite mixture models.Austin Bulletin,Volume 44,issue 02,pp 417-444 (2014).*

[TSF] Tzougas,G.,S. Vrontos and N. Frangos:*bonus-malus systems with two component mixture models arising from different parametric families (2018).*