ATHENS UNIVERSITY
OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

A study of optimal bonus-malus systems in automobile insurance using different underlying approaches

By

George J. Tzougas

A THESIS

Submitted to the Department of Statistics

of the Athens University of Economics and Business

in partial fulfilment of the requirements for

the degree of Master of Science in Statistics

Athens, Greece

July 2008

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Μελέτη των βέλτιστων συστημάτων εκπτώσεων-επιβαρύνσεων στην ασφάλιση οχημάτων χρησιμοποιώντας διαφορετικές βασικές προσεγγίσεις

Γεώργιος Ι. Τζουγάς

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Ιούνιος 2008

DEDICATION

To my parents Ioannis and Constantina

and my beloved family

ACKNOWLEDGEMENTS

I wish to express profound gratitude to my supervisor Professor Nicholas E. Frangos for his excellent research guidance, constant source of inspiration, encouragement and advice during this research work and my Masters program of study at the Department of Statistics of the Athens University of Economics and Business. Finally, I express my special gratitude and heartfelt thanks to both my father for his continuous support and my mother for her encouragement to go on with my education, this work is dedicated to them.

VITA

I was born in Athens on April 19, 1983. During my early school years, I found great interest in mathematics participating in several mathematical competitions and gaining awards. Also, at that time I took music lessons and played basketball for a local team. In July 2001, I graduated from Doukas High School with an average grade 17.6 and in the September of the same year I was enrolled, as a first year student, in the Department of Mathematics of the University of Athens, where I met most of my current friends. During my graduate studies, I gained my first working experience between 2004 and 2005 when I had full time active participation in several stages of the project ''A risk analysis information system'', supported by the General Secretary of Development and Investment, of the Ministry of National Economy and Finance. In 2006, I finished my graduate studies and I was awarded my bachelor degree in June 2007 with an average grade 7.4. Currently, during my postgraduate studies in the Department of Statistics of the Athens University of Economics and Business, I have successfully attained and finished with an average grade 7.2 the following courses: Theoretical Statistics I, Introduction to Statistical Packages and Data Analysis, Sampling Theory, Computational Statistics I, Time series, Generalized Linear Models, Stochastic Process with Emphasis in Finance I, Theoretical Statistics II, Computational Statistics II, Stochastic Process with Emphasis in Finance II. Furthermore, I have participated in several projects of the above courses and in one project in Multivariate Analysis. Finally, apart from Greek, I speak English (First Certificate in English, Certificate in Advanced English) and French and I have a good knowledge of the following software packages: MATLAB, R/S-Plus, E-Views, SPSS, Mathematica and C^{++}.

ABSTRACT

The parallel growth of accidents and casualties to the increasing number of motor vehicles during the twentieth century and up to our days, has led the actuarial scientists around the world to develop Bonus-Malus Systems (BMS) that penalize insureds responsible for one or more accidents by premium surcharges or *maluses* and reward claim-free policyholders by awarding them discounts or *bonuses*. The Bonus-Malus Systems have played a fundamental role in the automobile insurance since it holds a significant part of the non-life business of many companies. Furthermore, due to the enormous and still growing competitiveness of the market, the Bonus-Malus Systems should be efficient, penalizing the bad drivers and simultaneously competitive. A basic interest of the actuarial literature is the construction of an optimal or 'ideal' BMS defined as a system obtained through Bayesian analysis. The main objective of the current thesis will be the study of optimal Bonus-Malus Systems using different underlying approaches. The majority of optimal BMS assign to each policyholder a premium based on his number of claims (claim frequency) disregarding his/hers aggregate claim amount (claim severity). In this way, a policyholder who underwent an accident with a small size of loss will be unfairly penalized in the same way with a policyholder who had an accident with a big size of loss. Motivated by this, in chapter 2 we will present an optimal BMS based on the a posteriori frequency and the a posteriori severity component under the assumption that the number of claims is distributed according to the Negative Binomial distribution and that the losses of the claims are distributed according to the Pareto distribution. Also, we will present a generalized optimal BMS that is based both on the a priori and a posteriori classification criteria by incorporating a priori information for each policyholder in the above design. In chapter 3, we will present a classical optimal BMS that takes into account the claim frequency and one that takes into account both the claim frequency and the claim severity. This time the claim frequency is distributed according to the Geometric distribution and the claim severity is distributed according to the Pareto distribution again. In chapter 4 we will present an optimal BMS that uses a three parameters distribution the Hofmann's distribution for modeling claim frequency. Furthermore, a non-parametric method, that permits a simple formulation of the stationary and transition probabilities in a portfolio, is presented for the construction

of an optimal BMS. In chapter 5, our analysis is based on the fact that for the construction of optimal BMS the distribution of the number of claims is frequently chosen within the "mixed-Poisson" family. We will show the general properties of "mixed Poisson" family distributions and we will give a unifying approach of several particular cases including the geometrical, the P-Erlang, the Negative Binomial and the Poisson inverse gaussian distributions. Also, in order to avoid the problem of adjustment that is the thickness of the tails of the underlying distributions we will present a new family of "mixed-Poisson", built upon "fatty-tailed" underlying distributions, the "P-rational'' distributions. In chapter 6, we will present an alternative approach to BMS the Stochastic Vortices Model developed under the assumption that we have an open portfolio, i.e., we consider that a policy can be transferred from one insurance company to another and that the new policies that constantly arrive into a portfolio can be placed not only in the "starting class" but into any of the bonus classes. The Stochastic Vortices Model applies to populations divided into sub-populations which correspond to the transient states of homogeneous Markov chains. Also, by using the limit state probabilities of the Model we can estimate the Long Run Distribution for a BMS and calculate optimal bonus-malus scales. Furthermore, since the Stochastic Vortices Model allows the subscription and the annulment of policies in the portfolio it is an alternative approach to the usual BMS model and the fact that the population is taken as open renders it quite representative of the reality. Finally, in chapter 7 for the first time in actuarial literature, we will propose a combination of a Poisson- Inverse Gaussian distribution for modeling claim frequency and of a Pareto distribution for modeling claim severity for the construction of an optimal BMS.

ΠΕΡΙΛΗΨΗ

Η παράλληλη αύξηση των ατυχημάτων και των απωλειών ζωής με τον αυξανόμενο αριθμό μηχανοκίνητων οχημάτων κατά τη διάρκεια του εικοστού αιώνα μέχρι και τις μέρες μας, έχει οδηγήσει τους επιστήμονες του αναλογισμού ανά τον κόσμο να αναπτύξουν συστήματα εκπτώσεων- επιβαρύνσεων (Bonus-Malus Systems) τα οποία επιβάλουν ποινές στους ασφαλιζομένους που είναι υπαίτιοι για ένα ή περισσότερα ατυχήματα μέσω της επιβολής επιβαρύνσεων στα ασφάλιστρά τους (ή malus) και επιβραβεύουν τους ασφαλιζομένους χωρίς ατυχήματα απονέμοντας τους εκπτώσεις (ή bonus). Τα συστήματα εκπτώσεων- επιβαρύνσεων έχουν παίξει έναν θεμελιώδες ρόλο στην ασφάλιση των οχημάτων καθώς κατέχουν ένα σημαντικό τμήμα στον κλάδο γενικών ασφαλίσεων (non-life business) πολλών εταιρειών. Επιπλέον, λόγω της ήδη τεράστιας αλλά και επιπλέον αυξανόμενης ανταγωνιστικότητας της αγοράς, τα συστήματα εκπτώσεων- επιβαρύνσεων πρέπει να είναι συνεπή επιβάλλοντας ποινές στους κακούς οδηγούς και ταυτόχρονα ανταγωνίστηκα. Βασικό ενδιαφέρον της αναλογιστικής επιστήμης αποτελεί η κατασκευή ενός βέλτιστου ή ‘ιδανικού’ συστήματος εκπτώσεων- επιβαρύνσεων το οποίο ορίζεται ως ένα σύστημα το οποίο αποκτήθηκε μέσω της Μπευζιανής ανάλυσης. Ο βασικός στόχος της παρούσας διατριβής είναι η μελέτη των βέλτιστων συστημάτων εκπτώσεων- επιβαρύνσεων χρησιμοποιώντας διαφορετικές βασικές προσεγγίσεις. Η πλειονότητα των βέλτιστων συστημάτων εκπτώσεων- επιβαρύνσεων βασίζεται στον αριθμό των ατυχημάτων (συχνότητα των ατυχημάτων) του ασφαλιζόμενου για να την ανάθεση του ασφάλιστρου και αγνοεί το συνολικό τους κόστος (σφοδρότητα των ατυχημάτων). Συνεπώς αδίκως η επιβαλλόμενη ποινή σε ένα ασφαλιζόμενο ο οποίος υπέστη ατύχημα μικρού κόστους θα είναι όμοια με ένα ασφαλιζόμενο ο οποίος υπέστη ατύχημα μεγάλου κόστους. Παρακινούμενοι από αυτό το γεγονός στο κεφάλαιο 2, υπό την παραδοχή ότι ο αριθμός των ατυχημάτων κατανέμεται με βάση την Αρνητική Διωνυμική κατανομή και ότι το κόστος των ατυχημάτων κατανέμεται με βάση την κατανομή Pareto, θα παρουσιάσουμε ένα βέλτιστο σύστημα εκπτώσεων- επιβαρύνσεων το οποίο βασίζεται στην a posteriori συνιστώσα της συχνότητας αλλά και στην a posteriori συνιστώσα της σφοδρότητας των ατυχημάτων. Επιπλέον, θα παρουσιάσουμε ένα γενικευμένο βέλτιστο σύστημα εκπτώσεων- επιβαρύνσεων το οποίο βασίζεται και στα a priori και στα a posteriori κριτήρια κατηγοριοποίησης των

ασφαλιζομένων καθώς ενσωματώνει στο αρχικό σύστημα και τις a priori πληροφορίες για κάθε ασφαλιζόμενο. Στο κεφάλαιο 3 θα παρουσιάσουμε ένα κλασικό βέλτιστο σύστημα εκπτώσεων- επιβαρύνσεων, το οποίο λαμβάνει υπόψη την συχνότητα των ατυχημάτων, και ένα άλλο το οποίο λαμβάνει υπόψη και την συχνότητα των ατυχημάτων και την σφοδρότητα των ατυχημάτων. Αυτή τη φορά η συχνότητα των ατυχημάτων θα κατανέμεται με βάση την Γεωμετρική κατανομή και η σφοδρότητα των ατυχημάτων θα κατανέμεται, πάλι, με βάση την κατανομή Pareto. Στο κεφάλαιο 4 θα παρουσιάσουμε ένα βέλτιστο σύστημα εκπτώσεων-επιβαρύνσεων στο οποίο χρησιμοποιείται μια κατανομή τριών παραμέτρων, η κατανομή Hofmann, για να μοντελοποιηθεί η συχνότητα των ατυχημάτων. Επιπλέον θα παρουσιάσουμε μια απαραμετρική μέθοδο η οποία ‘επιτρέπει’ μια απλή τροποποίηση των στάσιμων (stationary) πιθανοτήτων και των μεταβατικών (transition) πιθανοτήτων ενός χαρτοφυλακίου για την κατασκευή ενός βέλτιστου συστήματος εκπτώσεων- επιβαρύνσεων. Στο κεφάλαιο 5 η ανάλυσή μας θα βασιστεί στο γεγονός ότι για την κατασκευή ενός βέλτιστου συστήματος εκπτώσεων-επιβαρύνσεων η κατανομή του αριθμού των ατυχημάτων επιλέγεται συχνότερα μέσα από την οικογένεια μεικτών κατανομών Poisson (“mixed Poisson” family). Θα δείξουμε τις γενικές ιδιότητες των κατανομών από την οικογένεια μεικτών κατανομών Poisson και θα δώσουμε μια ενοποιημένη προσέγγιση αρκετών συγκεκριμένων κατανομών όπως της Γεωμετρικής, της P-Erlang, της Αρνητική Διωνυμικής και της Poisson inverse Gaussian. Επίσης αποσκοπώντας στην αποφυγή του προβλήματος της προσαρμογής το οποίο αφορά το πάχος των ουρών των ανωτέρω βασικών κατανομών θα παρουσιάσουμε μια καινούργια οικογένεια μεικτών κατανομών Poisson η οποία έχει δημιουργηθεί βάση κατανομών με ‘παχιές’ ουρές τις “P-rational’’ κατανομές. Στο κεφάλαιο 6 θα παρουσιάσουμε μια διαφορετική προσέγγιση των συστημάτων εκπτώσεων- επιβαρύνσεων το Stochastic Vortices Model το οποίο έχει δημιουργηθεί υποθέτοντας ότι έχουμε ανοιχτό χαρτοφυλάκιο, δηλαδή θεωρούμε ότι ένα ασφαλιστήριο συμβόλαιο επιτρέπεται να μεταφερθεί από την μια ασφαλιστική εταιρεία στην άλλη και ότι τα νέα ασφαλιστήρια συμβόλαια που συνεχώς καταφθάνουν σε ένα χαρτοφυλάκιο δεν τοποθετούνται μόνο στην αρχική κατηγορία αλλά σε οποιαδήποτε κατηγορία bonus . Το Stochastic Vortices Model εφαρμόζεται σε πληθυσμούς που χωρίζονται σε υποπληθυσμούς οι οποίοι αντιστοιχούν στις μεταβατικές καταστάσεις των ομογενών Μαρκοβιανών αλυσίδων. Επίσης χρησιμοποιώντας τις οριακές πιθανότητες καταστάσεων (limit state probabilities) του μοντέλου μπορούμε να εκτιμήσουμε την Long Run Distribution ενός μπόνους-μάλους συστήματος και να υπολογίσουμε βέλτιστες bonus-malus κλίμακες. Επιπλέον το Stochastic Vortices Model αποτελεί μια διαφορετική προσέγγιση των συστημάτων εκπτώσεων- επιβαρύνσεων επειδή επιτρέπει την συνδρομή αλλά και την κατάργηση των ασφαλιστηρίων συμβολαίων στο χαρτοφυλάκιο. Εν κατακλείδι στο κεφάλαιο 7 για πρώτη φορά στην βιβλιογραφία της ασφαλιστικής επιστήμης θα προτείνουμε τον συνδυασμό της Poisson- Inverse Gaussian κατανομής, για την μοντελοποίηση της συχνότητας των ατυχημάτων, με την κατανομή Pareto για την μοντελοποίηση της σφοδρότητας των ατυχημάτων με σκοπό την κατασκευή ενός βέλτιστου συστήματος εκπτώσεων- επιβαρύνσεων.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Automobile insurance

During the previous century and up to our days, the number of motor vehicles is under a constant growth. Unfortunately, a bad consequence of this event is the parallel growth of accidents and casualties.

Property and liability motor vehicle coverage is divided into first and third party coverage. First party coverage provides protection if the vehicle's owner is responsible for the accident and protects him as well as his property. Third party coverage provides protection if the vehicle's owner causes harm to another party, who recovers their cost from the policyholder. First party coverage may include first party injury benefits such as medical expenses, death payments and comprehensive coverages.

Automobile third-party liability coverage has been made compulsory in most developed countries for a vehicle to be allowed on the public road network. The compulsory motor third party liability coverage represents a considerable share of the yearly nonlife premium collection in these developed countries. Furthermore, this share becomes more prominent when first party coverages are considered (i.e. medical benefits, uninsured or underinsured motorist coverage, and collision and other than collision insurance). Moreover, insurance companies maintain large data bases, recording policyholders' characteristics as well as claim histories. The economic importance and the availability of detailed information have to do with the fact that a large body of the actuarial literature is devoted to this line of business. Most actuaries from all over the world are devoted to the analysis of the number of claims filed by an insured driver over time. They face the problem of designing tariff structures that will fairly distribute the burden of claims among policyholders.

The formation of a 'pool' in which the policyholders put their risks is the fundamental principle of insurance. Also, it is fair to ask each member to pay a premium that is proportional to the risk that he imposes on the pool if these risks are unequal. In order to have a fair share of the costs of claims, it is very important to estimate the underlying risk of each insured party for the construction of a set of rates. In the design of a new tariff structure, the main task of the actuary is to make this possible by the partition of all policies into homogeneous classes with all policyholders paying the same premium when they belong to the same class.

Policyholders are covered by the liability insurance if, as drivers of a covered vehicle, injure a third party's property. The insurer provides legal defense for the policyholder if the policyholder is sued with respect to negligence for property damage or bodily injury. Also, the damages assessed against the policyholder will be paid by the insurer on behalf of the insured.

In the tort system, the claim is indemnified by the insurer only if he believes that the accident was caused by the insured or the third party sues the insured and proves that he was at fault in the accident. In such a system a large part of the premium income is consumed by legal fees, court costs and insurers' administration expenses. That is the reason why, several North-American jurisdictions have implemented a no-fault motor insurance system.

Nevertheless, even in a pure no-fault motor environment, the police still investigates which driver caused the accident or the degrees to which the drivers shared the fault because the at-fault events cause the insurance premium to rise at the next policy renewal.

A key actuarial ratemaking principle is cost-based pricing of individual risks. An estimate of the future costs related to the insurance coverage is the price charged to policyholders. The price of an insurance policy is defined by the pure premium approach as the ratio of all future claims estimated costs against the coverage that the insurance policy provides while it is in effect to the risk exposure, plus expenses.

The ratemaking of the property/casualty is based both on a claim frequency distribution and a loss distribution. The claim frequency is defined as the number of incurred claims per unit of earned exposure. The exposure is measured in car-year for motor third party liability insurance, the rate manual lists rates per car-year. The average payment per incurred claim is the average loss severity. Under mild conditions, the pure premium equals the product of the average claim frequency times the average loss severity.

In liability insurance, several years are often required for the settlement of larger claims. Therefore much of the data available for the recent accident years will be incomplete, as the final claim cost will be unknown. In this case, a final cost estimate can be obtained by loss development factors. Then the average loss severity is based on incurred loss data. In contrast to paid loss data, which are purely objective representing the company's actual payments, incurred loss data include subjective reserve estimates. The large claims have to be analyzed carefully by the actuary due to the fact they represent a considerable share of the insurer's yearly expenses.

1.2 Risk classification

Nowadays, in a competitive market it is very difficult for insurance companies to maintain cross subsidies between different risk categories. For example if male drivers are proved to cause fewer accidents than female drivers and if a company charged an average premium to all policyholders regardless of gender, disregarding this variable, most of its male policyholders would be tempted to move to another company that offers better rates to male drivers. Then a disproportionate number of female policyholders and insufficient premium income to pay for the claims are the result for the former company.

In order avoid lapses in a competitive market, as we have already mentioned, actuaries have to design a tariff structure that will distribute the burden of claims among policyholders fairly. The policies are partitioned into classes with all policyholders belonging to the same class paying the same premium. Every time an additional rating factor is used by a competitor, the actuary must refine the partition to avoid losing the best drivers with respect to this factor. Because of this, we can understand why insurance companies use so many factors even thought this is not required by actuarial theory, but instead is required by competition among insurers.

In a free market, a rating structure that matches the premiums for the risks as closely as possible or at least as closely as the rating structures used by competitors has to be used by insurance companies. This entails that virtually every available classification variable correlated to the risk must be used. If this doesn't happen the chance to select against competitors would be sacrificed, and the risk of suffering adverse selection by them is incurred. Thus the competition between insurers doesn't lead to actuarial science since it leads to more and more partitioned portfolios. Also, social disasters are caused very often by this trend towards more risk classification. More specifically, drivers sharing the characteristics of 'bad' drivers are tempted to drive without insurance as they do not find coverage for a reasonable price. At this point we should mention that even if a correlation exists between the rating factor and the risk covered by the insurer, there may be no causal relationship between that factor and risk. Requiring that insurance companies establish such a causal relationship to be allowed to use a rating factor is subject to debate.

Classification plans for the creation of risk classes are used by property and liability motor vehicle insurers. The classification variables introduced to partition risks into cells are called *a priori variables* because we can determine their values before the policyholder starts to drive. Premiums for motor liability coverage often vary by the use of the vehicle i.e. business use or driving to and from work, the territory in which the vehicle resides, and individual characteristics such as age, gender, occupation and marital status of the policyholder, for instance, if not precluded by legislation or regulatory rules. If any of these classification variables are misrepresented by the policyholders in their declaration, they can lose the coverage when they are involved in an accident. There is thus a strong incentive for accurate reporting of risk characteristics.

A priori classification can be easily achieved with the use of generalized regression models. The method can be summarized as follows: As the base cell we choose one risk classification cell. Normally it has the largest amount of exposure. The rate for the base cell is referred to as the base rate. A variety of risk classification variables, such as territory and so on define other rate cells. For each risk classification variable, there is a vector of differentials; with the base cell characteristics always assigned one hundred per cent.

At this point, it is worth mentioning that the generalized linear models developed by Nelder and Wedderburn (1972) can be very useful. Nelder and Wedderburn discovered that regression models with a response distribution belonging to the exponential family of probability distributions shared the same characteristics. Members of this family include distributions that have been widely used by actuaries to model the claims number, or claims severities such as Normal, Binomial, Poisson, Gamma and Inverse Gaussian. Actuaries who work with the exponential family can relax the restrictive hypotheses behind the Normal linear regression model, namely: the response variable takes on the theoretical shape of a Normal distribution, the variance is constant over individuals, the fitted values are obtained from linear predictors, or scores which are linear combinations of the explanatory variables.

More specifically, another member of the exponential family can replace the Normal distribution, this can be allowed for heteroscedasticity, and we can obtain fitted values from the link function a nonlinear transformation of linear predictors. Finally, the regression parameters can be estimated by maximum likelihood (many statistical packages offer efficient algorithms for this purpose).

1.3 Experience Rating

The supervising authorities excluded from the tariff structure certain risk factors, even though they were significantly correlated to losses, because of the trend towards more classification factors. Classifications based on items that are beyond the control of the insured such as gender or age, were banned by many states. The resulting inadequacies of the *a priori* rating system can be corrected for by the idea that came in the mid-1950s to use the past number of claims in order to allow for premiums adjustments *a posteriori*. Such practices are called *experience rating*. Furthermore, these practices are much in line with the concept of fairness: *a priori* ratemaking penalizes policyholders who 'look like' bad drivers, even if they are actually very good drivers and they will never cause any accident, whereas experience rating adjusts the amount of premium using the individual claim record. A balance between the likelihood of being a good driver, who is unlucky and suffered a claim, and the likelihood of being a truly bad driver, to whom his insurance company should make an increase in the premium he has to pay, is made by the use of actuarial credibility models. Also, experience rating system may be more acceptable to policyholders than seemingly arbitrary *a priori* classifications as it fair to correct the inadequacies of the *a priori* system by using a more adequate system.

Additionally, in the *a priori* risk classification many important factors cannot be taken into account even if the considerable importance of these factors is acknowledged by common sense and experience. More specifically, individual driving abilities such as accuracy of judgment, swiftness of reflexes, aggressiveness behind the wheel, respect for the highway code, and drinking behavior, are not taken into account in auto insurance rating, *a priori,* as these variables are impossible to measure in a cost-efficient way. For instance, consider two middle-aged males who are driving the same car model in the same city. Due to differences in individual behavior, these drivers may exhibit very different accident patterns. Consequently, tariff cells are still quite heterogeneous (we can model this heterogeneity by a random effect in a statistical model). So, it seems only reasonable that the number of claims reported by the policyholders can partly reveal hidden characteristics. The adjustment of the premium from the individual claims history can restore fairness among policyholders. Hence, an exogeneous explanation of serial correlation for longitudinal data is responsible for the allowance of past claims in a rating model. Then, correlation is only apparent and derives from the revelation of hidden features in the risk characteristics.

At this point, it should be mentioned that an endogeneous explanation can also be received by serial correlation for claim numbers. Accordingly, the claims history of the policyholders modifies the risk they represent. For example, it is possible that when a claim is reported, the risk of reporting a future claim may lower as the perception of danger while driving may be modified. Thus incentives to careful driving are provided by experience rating schemes and also negative contagion should induce. Nevertheless, since we always observe positive contagion for the claims number the main interpretation for automobile insurance is exogeneous. That is because policyholders who reported claims in the past have a bigger probability to produce claims in the future than those who did not, whereas true contagion should be negative.

1.4 Bonus-Malus Systems in automobile insurance

Experience rating is used by insurers in many European and Asian countries, as well as in North-American states. Such systems penalize insureds responsible for one or more accidents by premium surcharges or *maluses* and reward claim-free policyholders by awarding them discounts or *bonuses*. Their main purpose except encouraging policyholders to drive carefully is to better assess individual risks so that in the long run, everyone will pay a premium that corresponds to his own claim frequency. Depending on the country, insurers will refer to these systems as: no-claim discounts, experience-rated, personalized premiums, merit-rating systems, driving penalty points or bonus-malus systems (BMS).

In the United Kingdom, discounts for claim-free driving have been awarded very early at 1910 but their intention at that time was to induce the renewal of a policy with the same company and not to reward careful driving. Grenander (1957a,b) through his pioneering works was the first to provide theoretical treatments of bonus-malus systems. The first ASTIN colloquium was held in France in 1959 and it was exclusively devoted to no-claim discounts in insurance, with particular reference to motor business.

Many countries around the world use various bonus-malus systems. A typical form of no-claim bonus in the United Kingdom is defined as follows: An extra year of bonus is earned by drivers for each year they remain without claims at fault up to a maximum of four years, but two years bonus is lost each time they report a claim at fault. In such a system, maximum bonus is achieved in only a few years and the majority of mature drivers have maximum bonus.

In Continental Europe the bonus-malus systems that are used are often more elaborate. Bonus-malus scales consist of a finite number of levels, each with its own relativity or relative premium. Then the amount of premium that is paid by a policyholder is the product of a base premium with the relativity corresponding to the level occupied in the scale. New policyholders enter a specified level and according to transition rules of the bonus-malus system, after each year the policy moves up or down. If a bonus-malus system is in force, all policies in the same tariff class are partitioned according to the level they occupy in the bonus-malus scale. Hence, the bonus-malus systems can be considered as a refinement of *a priori* risk evaluation as according to individual past claims histories, they split each risk class into a number of subcategories.

The mathematical definition of a bonus-malus system requires the system to be Markovian., bonus-malus systems can be modelled using conditional Markov chains, provided they possess a certain ‘‘memory-less’’ property: the knowledge of the present class and of the number of claims of the present year suffices to determine the class for the next year. More precisely, the Markov property is satisfied by the bonus-malus systems as follows: the future level of year t+1, depends on the present level of year t and the number of accidents reported during that year and does not depend on the past i.e. the claim history and the levels occupied during years 1,2,... ,t— 1. Thus, we can determine the optimal relativities using an asymptotic criterion based on the stationary distribution or using transient distributions. Nevertheless, the Markov property is not satisfied for several BMS in force. Fortunately, in these cases a subdivision of the bonus-malus classes can be used to analyze the systems within the framework of the theory of finite Markov chains.

During the 20th century, a uniform bonus-malus system was imposed on all the companies in most of the European countries. In 1994, the European Union directed that the mandatory bonus-malus systems must be dropped by its entire member countries because they were in contradiction to the total rating freedom implemented by the Third Directive and as well the competition between insurers was reduced. Since then, Belgium for instance, dropped its mandatory system but the former uniform system is still applied by all companies operating there with minor modifications for the policyholders who occupy the lowest levels in the scale. However, in other European countries, like Spain and Portugal, insurers compete on the basis of bonus-malus systems. Nevertheless, the mandatory systems in France and Grand Duchy of Luxembourg are still in force as in 2004 the European Court of Justice decided that both these mandatory systems were not contrary to the rating freedom imposed by the European legislation. Thus, the French law still imposes on the insurers operating in France a unique bonus-malus system which is not based on a scale but instead uses the concept of an increase-decrease coefficient: a malus of 25 %

per claim and a bonus of 5 % per claim-free year are implied, so a base premium is assigned to each policyholder and it is adapted according to the claims number reported to the insurer, multiplying by 1.25 the premium each time an accident at fault is reported, and by 0.95 per claim-free year.

1.4.1 Tools for the evaluation of different bonus-malus systems

Due to the fact that regulatory environments in many countries are extremely diversified, from total freedom to government-imposed systems, with many intermediate situations, actuaries have developed tools for the evaluation of the different bonus-malus systems in force around the world. Also, as the approach to bonus-malus design depends on regulation more sophisticated BMS are expected in heavily regulated countries. Furthermore, the comparison of BMS across countries may prove to be interesting, because it allows insurers to evaluate how "severe" their BMS is, compared to neighbors.

The main goal of the evaluation of BMS is to rank all systems according to an *Index of Toughness.* Obviously, tough is not to be considered as a synonym of good or bad and the ranking of all BMS according to *their toughness* does not imply any judgment about their quality. Also, in all fairness to the different BMS, it must be noted that they must be compared under the same hypotheses. For instance, Lemaire(1994) compares thirty different systems, from twenty-two countries simulated, under the Poisson assumption.

Four important measures of toughness that can be used for the comparison of the different BMS across countries are the following:

- The *relative stationary average level* (RSAL).

  The RSAL is a measure of the degree *of* clustering of the policies in the lowest classes *of* a BMS. It measures the position of the average driver once the BMS has reached the steady-state condition as it relates the average simulated premium level, once the BMS has reached the steady-state condition, to the minimum and maximum levels. An evaluation of the treatment of new policyholders is enabled by a simple comparison of the RSAL with the starting level.

- The *coefficient of variation of the insured's premiums*.

  Insurance is practically a transfer of risk from the insured to the insurer. This transfer is total, perfect solidarity, without experience rating. With experience rating, according to the claims history of the policyholders, the personalized premiums will vary from year to year and solidarity between policyholders is weakened. Nevertheless, the variability of payments inherent in all experience rating systems has been used by their critics who argue that it goes against economic stability. The coefficient of variation (standard deviation divided by mean) measures this variability of annual premiums and can be used to evaluate the solidarity between the insureds. Also, because the coefficient of variation is a dimensionless parameter there is no need for currency conversions.

- The *elasticity* of BMS.

  The fact that almost all BMS in force around the world penalize the number of claims, independently of their amount, has the important implication that the risk of each driver can be measured by his individual claim frequency λ.It is an indication that regulators and insurance companies of most countries seem to accept it, at least as a good approximation.

  The elasticity of BMS is a measure of the response of the system to a change in the number of claims. For any reasonable BMS, the lifetime premiums paid by policyholders should be an increasing function of λ. Finally, we should mention that there are two versions of this measure: The asymptotic concept that requires the calculation of the stationary probability class distribution of the Markov chain and the transient concept that requires the calculation of the discounted expectation of lifetime payments.

- The *average optimal retention*.

  Since the introduction of early BMS, a very well known side-effect of this form of rating was the tendency of policyholders to pay small claims themselves and not report them to their carrier in order to avoid future premium increases.

  The optimal strategy of the insured -the limiting claim amount, the retention, under which he should indemnify his victims, and above which he should report the accident to his company- depends on several factors. The most important of these factors are his present BMS class, his claim frequency, and his discount rate. Lemaire (1994) provides as with an algorithm to compute optimal retentions and introduces a currency conversion factor in order to enable the comparison of the average optimal retention across countries.

The four measures of toughness mentioned above are positively correlated. A severe BMS that penalizes claims heavily will exhibit high *relative stationary average level,* high *coefficient of variation of premiums,* high *elasticity,* and high *average optimal retentions.* Furthermore, factor analysis can be used to summarize the data and define a single measure of toughness. By the use of a single principal component most of the information contained in the data can be retained. A summary measure for the evaluation of the severity of all systems an *Index of Toughness* can be defined by scores and this first component. Also, by assigning standardized scores to each BMS, all systems can be ranked according toughness.

1.4.2 Actuarial and economic justifications for bonus-malus systems

Another important issue is the actuarial and economic justifications for BMS.As we have already mentioned, BMS, by taking into consideration the claims history of each policyholder, allow premiums to be adapted for hidden individual risk factors and promote careful driving. This holds by asymmetrical information between the insurance company and the policyholders. It can be noticed that asymmetric information appears in insurance markets when it is difficult for firms to judge the riskiness of those who purchase insurance coverage and two aspects of this phenomenon are adverse selection and moral hazard. Adverse selection arises when the policyholders take advantage of the information about their claim behavior known to them but unknown to the insurer. Nevertheless, in the context of compulsory motor third party liability insurance, the problem of adverse selection is not as important as the problem of moral hazard when the insurance companies charge the same premium amount to every policyholder. In a deregulated environment with companies using many risk classification factors things are even worst and we can’t avoid adverse selection as very heterogeneous driving behaviors are observed among policyholders who share the same *a priori* variables. For every related coverages, such as comprehensive damages, adverse selection is always important.

According to Rotschild and Stiglitz consideration about adverse selection, individuals through the contract they choose may reveal their underlying risk, an important fact when setting an adequate tariff structure. When unobservable heterogeneity occurs a more comprehensive coverage is usually chosen by ‘risky’ agents and low risk insurance applicants select high deductibles.

Also, a comparison between the approaches of economists' and actuaries' to experience rating proves to be interesting. In order to counteract the efficiency which occurs from moral hazard, the economists introduce discounts and penalties while actuaries’ basic interest is to better assess the individual risk so that everyone will pay, in the long run, a premium corresponding to his own claim frequency. Thus, actuarial literature is more interested in adverse Section than moral hazard.

1.4.3 Optimal bonus-malus systems

A basic interest of the actuarial literature is the construction of an optimal or ‘ideal’ BMS defined as a system obtained through Bayesian analysis. The construction of an optimal BMS is a form of statistical games between the actuary and the nature. A BMS is called optimal if it is *financially balanced for the insurer*: the total amount of bonuses must be equal to the total amount of maluses and if it is *fair for the policyholder*: the premium paid for each policyholder is proportional to the risk that he imposes to the pool.

Furthermore, Optimal BMS can be divided in two categories:

- Those based only on the a posteriori classification criteria i.e. the number of claims and the severity of claims. They are function of time and of past number and amount of accidents
- Those based both on the a priori and the a posteriori classification criteria. This category, also takes into consideration the characteristics of each individual such as the age variable.

In both categories, the majority of the authors in actuarial literature designed BMS based on the number of accidents disregarding their severity. In the first category, BMS based on the a posteriori claim frequency component, we come across the work of Lemaire (1995) who designed an optimal BMS based on the number of claims of each policyholder, following the pioneer game-theoretic framework by Bichsel and Bühlmann (1964). Each policyholder must pay a premium proportional to his unknown claim frequency. The loss of the insurer derives from the use of the estimated claim frequency instead of the true unknown claim frequency and the estimate of the policy-holder's claim frequency that minimizes the loss incurred is the optimal one. Furthermore, Lemaire considered the optimal BMS obtained using the quadratic error loss function, the expected value premium calculation principle and the Negative Binomial as the claim frequency distribution. Tremblay (1992) designed an optimal BMS using the quadratic error loss function, the zero-utility premium calculation principle and the Poisson-Inverse Gaussian as the claim frequency distribution. Coene and Doray (1996) obtained a financially balanced BMS by minimizing a quadratic function of the difference between the premium for an optimal BMS with an infinite number of classes, weighted by the stationary probability of being in a certain class and by imposing various constraints on the system. Walhin and Paris (1997) obtained an optimal BMS using as the claim frequency distribution the Hofmann's distribution, which encompasses the Negative Binomial and the Poisson-Inverse Gaussian, and also using as a claim frequency distribution a finite Poisson mixture. In the second category, BMS based on the a priori and the a posteriori claim frequency component, we come across the work of: Dionne and Vanasse (1989, 1992) who designed a BMS that integrates a priori and a posteriori information on an individual basis. This BMS is a function of the years that the policyholder is in the portfolio, the claims number and the individual characteristics, which are significant for the number of accidents. So, as mentioned in Dionne and Vanasse (1989) by also taking into consideration the characteristics of each individual, the premiums vary simultaneously with other variables that affect the claim frequency distribution. For example, let us suppose that the age variable has a negative effect on the expected number of claims. This would imply that insurance premiums should decrease with age, even though premium tables derived from

BMS based only on the a posteriori criteria, do not allow for a variation of the statistically significant variable of age. Picech (1994) and Sigalotti (1994) developed a BMS that incorporates the a posteriori and the a priori classification criteria, with a single a priori rating variable, the engine power. Sigalotti derived a recursive procedure to compute the sequence of increasing equilibrium premiums needed to balance out premiums income and expenditures compensating for the premium decrease created by the BMS transition rules. Picech constructed a BMS that approximates the optimal merit-rating system using a heuristic method. Taylor (1997) developed a Bonus-Malus scale that uses some rating factors to recognize the differentiation of underlying claim frequency by experience, but only to the extent that this differentiation is not recognized within base premiums. Pinquet (1998) designed an optimal BMS from different types of claims, such as those at fault and those not at fault.

The optimal bonus-malus systems mentioned above, assign to each policyholder a premium based on the number of his accidents but the size of loss that each accident incurred is not considered. This is unfair, because a policyholder who underwent an accident with a small size of loss is penalized in the same way with a policyholder who had an accident with a big size of loss. Actually, this is something that Lemaire (1995) pointed out for the majority of bonus-malus systems in force throughout the world, which penalize the number of accidents without taking into account their severity. Korea is an exception because policyholders who had a bodily injury claim pay higher maluses, depending on how severe the accident was, than the policyholders who had a property damage claim.

Among the BMS that take severity into consideration are those designed from Picard (1976) and Pinquet (1997). Picard (1976) in order to take into account the subdivision of claims into two categories, small and large losses, generalized the Negative Binomial model and in order to separate large from small losses, used two options: i) A limiting amount under which the losses are regarded as small and the remainder as large. ii) More severe penalization of the policyholders who had a bodily injury accident than those that caused property damage in order to subdivide those two categories. Pinquet (1997) starting from a rating model based on the analysis of number of claims and of costs of claims and adding two heterogeneity components, to represent unobserved factors that explain the severity variables, designed an optimal BMS which makes allowance for the severity of the claims. Furthermore, i) the costs of claims were supposed to follow gamma or lognormal distribution, ii) the rating factors, as well as the heterogeneity components were included in the scale parameter of the distribution. ii) the heterogeneity was also considered to follow a gamma or lognormal distribution and a credibility expression, which provides a predictor for the average cost of claim for the following period was obtained.

1.5 Overview of the chapters 2 to 8

The study of optimal bonus-malus systems in automobile insurance using different underlying approaches will be the subject of interest of the current thesis. In chapter 2, following the work of Frangos and Vrontos (2001) we will present an optimal BMS based on the a posteriori frequency and the a posteriori severity component under the assumption that the number of claims is distributed according to the Negative Binomial distribution and that the losses of the claims are distributed according to the Pareto distribution. By the application of Bayes' theorem, we will find the posterior distributions of the mean claim frequency and the mean claim size, given the information we have about the claim frequency history and the claim size history for each policyholder for all the time that he stayed in the portfolio. Furthermore, important a priori information for each policyholder is incorporated in the above design of optimal BMS and a generalized BMS that is based both on the a priori and a posteriori classification criteria is proposed. In this generalized BMS the premium is a function of the years that the policyholder was in the portfolio, his number of accidents, the size of loss of each of these accidents and of the statistically significant a priori rating variables for the number of accidents and for the size of loss that each of these claims incurred.

In chapter 3, based on the work of Mert and Saykan (2005) we will present an optimal BMS that takes into account the claim frequency and one that takes into account both the claim frequency and the claim severity. This time for modeling claim numbers, we use a mixture of a Poisson and an Exponential distributions, the Geometric distribution and for modeling claim amounts we use the Pareto distribution again. The risk premium is calculated using the net premium principle, and the results are obtained by using the claim number only and by using both the claim number and claim amount.

In chapter 4 based on the work of Walhin and Paris (1997), which we have mentioned earlier, we will construct an optimal BMS using as the claim frequency distribution a three parameters distribution encompassing the Negative Binomial and the Poisson-Inverse Gaussian the Hofmann's distribution. Furthermore, a non-parametric method, that permits a simple formulation of the stationary and transition probabilities in a portfolio, is proposed for the construction of an optimal BMS.

In chapter 5, our analysis is based on the fact that for the construction of optimal BMS the distribution of the number of car accidents is frequently chosen within the "mixed-Poisson" family. Following the work of Hulin and Justens (Banque Pictet-Luxembourg HEFF/Cooremans Bruxelles) we will show the general properties of "mixed Poisson" family distributions and we will give a unifying approach of several particular cases including the geometrical, the P-Erlang, the Negative Binomial and the P-inverse gaussian distributions. Also, in order to avoid the problem of adjustment that is the thickness of the tails of the underlying distributions we will introduce a new family of "mixed-Poisson", built upon "fatty-tailed" underlying distributions, the "P-rational'' distributions.

In chapter 6, we will present an alternative approach to BMS the Stochastic Vortices Model based on the work of Guerreiro and Mexia (2004). We will develop the Stochastic Vortices Model under the assumption that we have an open portfolio, i.e., we consider that a policy can be transferred from one insurance company to another and that the new policies that constantly arrive into a portfolio can be placed not only in the “starting class” but into any of the bonus classes. The Stochastic Vortices Model applies to populations divided into sub-populations which correspond to the transient states of homogeneous Markov chains. Also, by using the limit state probabilities of the Model we can estimate the Long Run Distribution for a BMS and calculate optimal bonus-malus scales. Furthermore, since the Stochastic Vortices Model allows the subscription and the annulment of policies in the portfolio it is an alternative approach to the usual BMS model and the fact that the population is taken as open renders it quite representative of the reality.

In chapter 7 for the first time in actuarial literature, we will propose a combination of a Poisson- Inverse Gaussian distribution for modeling claim frequency and of a Pareto distribution for modeling claim severity for the construction of an optimal BMS. Using Bayes’ theorem we will calculate the posterior distributions of the mean claim frequency and the mean claim size, given the information we have about the claim frequency history and the claim size history for each insured. Optimality will be obtained by minimizing the insurer’s risk. At this point, we should mention that we will choose the Poisson- Inverse Gaussian distribution for modeling claim frequency as an alternative to the Negative Binomial distribution, that is frequently used, based on the fact that mixed Poisson distributions have thicker tails than the Poisson distribution thus they provide a better fit to claim frequency data than the Poisson distribution when the portfolio is heterogeneous. Furthermore, following Frangos and Vrontos (2001) and Mert and Saykan (2005) we will choose the heavy-tailed Pareto distribution for modeling claim severity because through the use of long tail distributions apart from many small claim severities, high claim severities can also be observed.

In chapter 8 we will present the conclusions of the current thesis. Finally, it should be mentioned that since it was practically impossible to summarize all the creative work of many researchers about the topic of optimal BMS we will attempt to give some justice to them at the references in the end of the thesis.

CHAPTER 2

DESIGN OF OPTIMAL BONUS-MALUS SYSTEMS WHERE THE CLAIM FREQUENCY DISTRIBUTION IS NEGATIVE BINOMIAL AND THE CLAIM SEVERITY DISTRIBUTION IS PARETO

2.1 Introduction

In chapter 2 we design an optimal BMS based both on the number of accidents of each policyholder and on the size of loss for each accident incurred. Optimality is obtained by minimizing the insurer's risk. Furthermore, the a priori information we have for each policyholder is incorporated in the above design of optimal BMS, so a generalized BMS is proposed that takes into consideration the individual's characteristics, his claims frequency and his claims severity.

2.2 Design of optimal BMS with a frequency and a severity component based on the a posteriori criteria

Under the basic assumption that the number of claims of each policyholder is independent from the severity of each claim, we propose that the number of claims is distributed according to the Negative Binomial distribution and that the losses of the claims are distributed according to the Pareto distribution. Using Bayes' theorem, we will find the posterior distributions of the mean claim frequency and the mean claim size, given the information we have about the claim frequency and claim size history for each policyholder. Optimality is obtained by minimizing the insurer's risk.

2.2.1 Frequency component

For the frequency component we use the same structure used by Lemaire (1995). Under the assumptions that the portfolio is heterogeneous and that all policyholders have constant but unequal underlying risks to have an accident, the number of claims k, given the parameter λ,is considered to follow the Poisson(λ) distribution,

$$p_{\lambda}(k \mid \lambda) = \frac{\lambda^{k}}{k!} e^{-\lambda}$$

κ= 0,1,2,3,... and λ > 0 and λ denotes the different underlying risk of each policyholder to have an accident. For the structure function, it is assumed that λ ~ *gamma* (α, τ) and λ has a probability density function of the form:

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

with *E(λ)*= α/τ and *Vα r* (λ) = α/τ^2. Then it can be proved that the unconditional distribution of the number of claims *k* will be the Negative Binomial (α,τ), with probability density function:

$$\text{P(k)=}\binom{k+\alpha-1}{k}\left(\frac{\tau}{1+\tau}\right)^{\alpha}\left(\frac{1}{1+\tau}\right)^{k}$$

With *E*(*k*) = α/τ and *Var* (*k*) = (α/t) (1 + 1/τ). A desirable property of this distribution is that its variance exceeds its mean, something which is common for all mixtures of Poisson distributions and allows us to deal with data that present overdispersion.

For $K = \sum_{i=1}^{t} k_i$ the total number of claims that a policyholder had in t years, with k_i the number of claims that the policyholder had in the year *i*, *i* = 1,...,*t*. using the Bayes' theorem we obtain the posterior structure function of λ for a policyholder or a group of policyholders with claim history k_1,....k_t,

$$u(\lambda \mid k_1, \ldots, k_t) = \frac{(\tau + t)^{K+\alpha} \cdot \lambda^{K+\alpha-1} \cdot e^{-(\tau+t)\lambda}}{\Gamma(\alpha + K)}$$

and that is the probability density function of a *gamma*(α + *K*,*t* + τ). Using the quadratic error loss function the optimal choice of λ_{t+1} for a policyholder with claim history k_1, ...k_t will be the mean of the posterior structure function, that is:

$$\lambda_{t+1}(k_1,...,k_t) = \frac{\alpha+K}{t+\tau} = \bar{\lambda}\left(\frac{\alpha+K}{\alpha+t\bar{\lambda}}\right), \bar{\lambda} = \frac{\alpha}{\tau} \qquad (1)$$

Finally, an update of the parameters of gamma, from α and τ to α + *K* and *t* + τ respectively, is necessary due to the occurrence of K accidents in t years. The gamma has the important property of the stability of the structure function as it is a conjugate family for the Poisson likelihood.

2.2.2 Severity component

Denote as *x* the size of the claim of each insured and as *y* the mean claim size for each insured and assume that the conditional distribution of the size of each claim given the mean claim size *x*|*y,* for each policyholder is the exponential distribution with parameter *y* with a probability density function given by:

$$\mathrm{f(x|y)} = \frac{1}{\mathrm{y}} \cdot e^{-\frac{x}{y}}$$

for *x* > 0 and *y* > 0. *E*(*x*| *y*)=*y* and *var*(*x y*) = y^2. The mean claim size *y* takes different values and is not the same for all the policyholders. Thus, it is considered that the prior distribution of *y* is Inverse Gamma with parameters *s* and *m* and probability density function given by (see
for example Hogg and Klugman (1984)):

$$g(y) = \frac{\frac{1}{m} \cdot e^{-\frac{m}{y}}}{\left(\frac{y}{m}\right)^{s+1} \cdot \Gamma(s)}$$

The expected value of the mean claim size y is :

$$E(y) = \frac{m}{s-1}$$

The unconditional distribution of the claim size x is:

$$\text{P(X=x)}=\int_0^\infty \text{f(x|y)} \cdot g(y)dy =$$
$$= s \cdot m^s \cdot (x+m)^{-s-1}$$

the probability density of the Pareto distribution with parameters s and m. Thus Pareto distribution can be generated in the following way: if $x|y \sim$ *Exponential*(y) and $y \sim$ Inverse *Gamma*(s,m) then $x \sim$ *Pareto*(s,m). So the relatively tame exponential distribution gets transformed in the heavy-tailed Pareto distribution, which is good for modeling the claim severity instead of the exponential distribution which is often inappropriate. Taking y distributed according to the Inverse Gamma, heterogeneity that characterizes the severity of the claims of different policyholders is incorporated in the model. Such a generation of the Pareto distribution can be found in Herzog (1996) and in other actuarial papers but Frangos and Vrontos (2001) are the first who used it in the design of an optimal BMS.

Then the posterior distribution g(y|$x_1, \ldots, x_k$) is calculated in order to design an optimal BMS that will take into account the size of loss of each claim. Denote for a policyholder that was in the portfolio for t years:

1. $K=\sum_{i=1}^{t} k_i$ the total number of claims that he had

in t years with k_i the number of claims that the policyholder had in the year i, $i = 1,\ldots,t$.

2. The vector $x_1, x_2, \ldots, x_k$ his claim size history and $\sum_{k=1}^{K} X_k$ his total claim amount.

Applying Bayes' theorem:

$$g(y \mid x_1, \ldots, x_k) = \frac{\dfrac{1}{\left(m + \sum_{k=1}^{K} x_k\right)} \cdot e^{-\frac{m + \sum_{k=1}^{K} x_k}{y}}}{\left(\dfrac{y}{m + \sum_{k=1}^{K} x_k}\right)^{K+s+1} \cdot \Gamma(K+s)}$$

the posterior distribution of the mean claim size y given the claims size history of the policyholder $x_1, \ldots, x_k$ and it is the Inverse Gamma($s + K$, m+$\sum_{k=1}^{K} x_k$). This update of the parameters of the Inverse Gamma from s and m to s+K *and* m+$\sum_{k=1}^{K} x_k$ is necessary due to the occurrence of K claims in t years with aggregate claim amount equal to $\sum_{k=1}^{K} x_k$. Also Inverse Gamma distribution is said to have the important property of being conjugate with the exponential likelihood. Finally, the mean of the posterior distribution of the mean claim size is:

$$E(x \mid y) = \frac{m + \sum_{k=1}^{K} x_k}{s + K - 1} \qquad (2)$$

and the predictive distribution of the size of the claim of each insured x will be also a member of the Pareto family.

2.2.3 Calculation of the Premium according the Net Premium Principle

For a specific group of policyholders who stayed in the portfolio for t years and had K claims with $\sum_{k=1}^{K} x_k$ total claim amount , the expected number of claims $\lambda_{t+1}(k_1, ..., k_t)$,was given by (1) and the expected claim severity $y_{t+1}(x_1, ..., x_K)$ by (2).Thus:

$$\Pr emium = \frac{\alpha + K}{t + \tau} \cdot \frac{m + \sum_{k=1}^{K} x_k}{s + K - 1} \qquad (3)$$

the net premium that must be paid from that specific group of policyholders and in order to find it we have to know:

1. α and τ , the parameters of the Negative Binomial distribution, (see Lemaire (1995) for their estimation)
2. *s* and *m* the parameters of the Pareto distribution (see Hogg and Klugman (1984) for their estimation)
3. the number of years *t* that the policyholder was in the portfolio,
4. his number of claims *K* and his total claim amount $\sum_{k=1}^{K} x_k$

2.2.4 Properties of the Optimal BMS with a frequency and a severity component

- The system is fair. Through the Bayes' theorem, each insured is informed both for the number and the loss of his claims the time that he is in the portfolio and thus the premium he pays is proportional to his claim frequency and his claim severity. *We use the exact loss x_k that is incurred from each claim in order to have a differentiation in the premium for policyholders with the same number of claims, not just a scaling with the average claim severity of the portfolio. Frangos and Vrontos (2001)*

- The system is financially balanced. Every year the average of all premiums collected remains constant and equal to

$$P = \frac{\alpha}{\tau}\frac{m}{s-1} \qquad (4)$$

- Same premium , equal to(4),for all policy holders in the beginning .
- Premium increases proportionally to the number and to the severity of the claims.
- Premium always decreases when no accidents occur.
- Phenomenon of bonus hunger will decrease and the estimate the actual claim frequency will be more accurate, because claims with small cost will be reported, as claim severity is taken into consideration.
- The introduction of the severity component is more crucial than the number of claims for the insurer since it determines the expenses of the insurer from the accidents and thus the premium that must be paid.
- The mean of severity is not always robust and can be affected by variation. So in practice a more robust estimator could be used. (i.e. cutting of the data, M-estimator).

2.3 Design of optimal BMS with a frequency and a severity component based both on the a priori and the a posteriori criteria

Dionne and Vanasse (1989, 1992) presented a BMS that integrates risk classification and experience rating based on the number of claims of each policy-holder. We extend this model proposing a generalized BMS that integrates a priori and a posteriori information on an individual basis based both on the frequency and the severity component. *One of the reasons for this development is that premiums should vary simultaneously with the variables that affect the distribution of the number of claims and the size of loss distribution. Frangos and Vrontos (2001).*

Denote as GBM_F the generalized BMS obtained when only the frequency component is used and as GBM_S the generalized BMS obtained when only the severity component is used The premiums of the generalized BMS will be derived using the following multiplicative tariff formula:

$$Premium = GBM_F * GBM_S \tag{5}$$

2.3.1 Frequency Component

We will obtain the GBM_F by using the same structure used by Dionne and Vanasse (1989, 1992). For an individual i with an experience of t periods assume that the number of his claims K_i^j for period j follows the Poisson distribution with parameter λ_i^j and K_i^j are independent. Then the expected number of claims of i for period j is denoted by λ_i^j and consider that it is a function of the vector $c_i^j = (c_{i,1}^j, \ldots, c_{i,h}^j)$ of h individual's characteristics, which represent different a priori rating variables. Specifically assume that $\lambda_i^j = \exp(c_i^j \beta^j)$, non-negativity of λ_i^j is implied from the exponential function, β^j is the vector of the coefficients. The probability specification becomes:

$$P\left(K_i^j = k\right) = \frac{e^{-\exp\left(c_i^j \beta^j\right)} \left(\exp\left(c_i^j \beta^j\right)\right)^k}{k!}$$

For the determination of the expected number of claims, it is assumed that enough information is provided by the h individual characteristics. Using maximum likelihood methods the vector of the coefficients β^j can be obtained, an application can be found in Haus-mann, Hall and Griliches (1984). However, a random variable ε_i has to be introduced into the regression component if it is assumed that the a priori rating variables do not contain all the significant information for the expected number of claims. Folowing Gourieroux, Montfort and Trognon (1984a), (1984b), it can be written:

$$\lambda_i^j = \exp(c_i^j \beta^j + \varepsilon_\iota) = \exp(c_i^j \beta^j) u_i$$

and $u_i = \exp(\varepsilon_i)$, for a random λ_i^j. If u_i follows a gamma distribution with $E(u_1)=1$ and $Var(u_i)=1/a$, it can be proved that this parameterization does not affect the results if there is a constant term in the regression and the probability specification becomes:

$$P\left(K_i^j = k\right) = \frac{\Gamma(k+\alpha)}{k!\Gamma(\alpha)} \left[\frac{\exp\left(c_i^j \beta^j\right)}{\alpha}\right]^k \left[1 + \frac{\exp\left(c_i^j \beta^j\right)}{\alpha}\right]^{-(k+\alpha)}$$

negative binomial with parameters α and $\exp(c_i^j \beta^j)$. $E(u_i)= 1$ is chosen in order to have $E(e_i)= 0$ and then

$$E\left(K_i^j\right) = \exp\left(c_i^j \beta^j\right), Var\left(K_i^j\right) = \exp\left(c_i^j \beta^j\right)\left[1 + \frac{\exp\left(c_i^j \beta^j\right)}{\alpha}\right]$$

See Lawless (1987), Gourieroux, Montfort and Trognon (1984a) and Gourieroux, Montfort and Trognon (1984b) for more on the Negative Binomial regression. The insurer using past information over the t periods of claim frequency and of known individual characteristics over the t+1 periods has to calculate the best estimator of the expected number of accidents at period $t + 1$,denoted as

$$\hat{\lambda}_i^{t+1}\left(K_i^1, \ldots, K_i^t; c_i^1, \ldots, c_i^{t+1}\right)$$

For a policyholder with claim history $K_i^1,\dots,K_i^t$ and $c_i^1,\dots,c_i^{t+1}$ characteristics using Bayes theorem the posterior structure function is gamma with updated parameters:

$$\alpha+\sum_{j=1}^{t}K_{i}^{j}\ ,\ \frac{a}{\exp(c_i^j\beta^j)}+\mathrm{t}$$

and using the quadratic loss function the optimal estimator is:

$$\hat{\lambda}_{t+1}(K_i^1,\dots,K_i^t;c_i^1,\dots,c_i^t)=$$

$$=\int_0^{\infty}\lambda_i^{t+1}(K_i^{t+1},u_i)\cdot f(\lambda_i^{t+1}\mid K_i^1,\dots,K_i^t;c_i^1,\dots,c_i^t)d\lambda_i^{t+1}=$$

$$=\frac{1}{t}\sum_{j=1}^{t}\exp\left(c_i^j\beta^j\right)\left[\frac{\alpha+\sum_{j=1}^{t}K_i^j}{\alpha+t\exp\left(c_i^j\beta^j\right)}\right]$$

<u>Notes:</u>

1. $\sum_{j=1}^{t}K_{i}^{j}$ is the total number of claims in t periods of policyholder i.

2. For t=0, $\hat{\lambda}_i^1\equiv\exp(c_i^j\beta^j)$ only a priori rating is used in the first period.

3. When a regression component is limited to a constant β_0,the univariate model without a regression component is obtained Lemaire (1995), Ferreira (1974).

Now we will deal with the generalized bonus-malus factor obtained when the severity component is used. It will be developed in the following way.

2.3.2 Severity Component

Here we obtain the generalized bonus-malus factor with the use of the severity component. For an individual i with an experience of t periods, assume that the number of his claims for period j is denoted as K_i^j , his total number of claims is denoted as K and the loss incurred from his claim k for the period j is denoted as $X_{i,\mathrm{K}}^j$.

Then, the vector $X_{i,1}, X_{i,2}, ..., X_{i,K},$ is the information we have for his claim size history and the total claim amount over the t periods in the portfolio for the policyholder will be $\sum_{k=1}^{K} X_{i,k}$ .It is assumed that $X_{i,\mathrm{K}}^j$ an exponential distribution with parameter y_i^j .The cost for the insurer, is different for each policyholder because the expected claim severity is not the same for everyone. Thus, it is fair for each policyholder to pay a premium proportional to his mean claim severity It is considered that the expected claim severity is a function of the vector of the h individual's characteristics, $d_i^j = (d_{i,1}^j, ..., d_{i,h}^j)$ which represent different a priori rating variables. More specifically it is assumed that $y_i^j = \exp(d_i^j \gamma^j)$ where y_i^j denotes the mean or the expected claim severity of a policyholder i in period j *and* γ is the vector of the coefficients. The probability specification becomes:

$$P(X_{i,\mathrm{K}}^j = x) = \frac{1}{\exp(d_i^j \gamma^j)} \cdot e^{-\frac{x}{\exp(d_i^j \gamma^j)}}$$

, non-negativity of y_i^j is implied from the exponential function.

It is assumed that enough information is provided by the h individual characteristics for determining the expected claim severity. However a random variable z_i has to be introduced into the regression component, if one assumes that the a priori rating variables do not contain all the significant information for the expected claim severity. Thus it can be written:

$$y_i^j = \exp(d_i^j \gamma^j + \varepsilon_t) = \exp(d_i^j \gamma^j) w_i$$

And w_i = exp(z_i), for a random y_i^j. If w_i follows an inverse gamma(s,s-1) distribution with E (w_i)=1 and $Var(w_i)$= $\frac{1}{s-2}$,s>2 then y_i^j follows an inverse gamma(s,(s-1) $\exp(c_i^j \gamma^j)$) and the probability specification for $X_{i,K}^j$ becomes:

$$P\left(X_{i,k}^j = x\right) = s \cdot \left[(s-1)\exp\left(d_i^j \gamma^j\right)\right]^s . \left(x + (s-1)\exp\left(d_i^j \gamma^j\right)\right)^{-s-1}$$

Pareto distribution with parameters s and (s-1) $\exp(c_i^j \gamma^j)$ .The above parameterization does not affect the results if there is a constant term in the regression. E (w_i)=1 is chosen in order to have $E(z_i)$ = 0. We also have:

$$E\left(X_{i,k}^j\right) = \exp\left(d_i^j \gamma^j\right), Var(X_{i,k}^j) = \frac{\left[(s-1)\exp\left(d_i^j \gamma^j\right)\right]^2}{s-1}\left(\frac{2}{s-2} - \frac{1}{s-1}\right)$$

The insurer using past information over the t periods of claim severity and of known individual characteristics over the t+1 periods has to calculate the best estimator of the expected amount of accidents at period t + 1,denoted as

$$\hat{y}_i^{t+1}(X_{i,1},...,\hat{X}_{i,K}; d_i^1,...,d_i^{t+1})$$

For a policyholder with claim sizes $X_{i,1},...,X_{i,K}$ in t periods and characteristics $d_i^1,...,d_i^{t+1}$ using Bayes theorem the posterior distribution of the mean claim severity is inverse gamma with updated parameters:

s+K and $(s-1)\exp(d_i^j \gamma^j) + \sum_{k=1}^{K} X_{i,k}$ and using the quadratic loss function the optimal estimator of the mean claim severity is:

$$\hat{y}_i^{t+1}(X_{i,1},..,X_{i,K};d_i^1,...,d_i^{t+1}) =$$

$$= y_i^{t+1}(X_i^{t+1},w_i)f(y_i^{t+1} \mid X_{i,1},..,X_{i,K};d_i^1,...,d_i^{t+1})dy_i^{t+1} =$$

$$= \int_0^\infty \frac{(s-1)\frac{1}{t}\sum_{j=1}^{t}\exp\left(d_i^j\gamma^j\right)+\sum_{k=1}^{K}X_{i,k}}{s+K-1}$$

Note: If t=0, only a priori rating is used in the first period, it is $\hat{y}_i^1=\exp(d_i^1\gamma^j)$.

2.3.3 Calculation of the Premiums of the generalized BMS

From (5):

*Premium = GBM$_F$ * GBM$_S$ =*

$$= \frac{1}{t}\sum_{j=1}^{t}\exp(c_i^j\beta^j)\left[\frac{\alpha+\sum_{j=1}^{t}K_i^j}{\alpha+t\exp(c_i^j\beta^j)}\right]\frac{(s-1)\frac{1}{t}\sum_{j=1}^{t}\exp(d_i^j\gamma^j)+\sum_{k=1}^{K}X_{i,k}}{s+K-1} \qquad (6)$$

The number of years t that the policyholder is in the portfolio, his total number of accidents in t years and his aggregate claim amount in t years must be known.

For the frequency component of the generalized BMS, the dispersion parameter α and the vector β i.e. the parameters of the negative binomial regression model, can be estimated with the maximum likelihood method.

For the severity component of the generalized BMS, s and γ^j can be estimated by using the quasi-likelihood and according to Renshaw (1994), who uses as a modeling tool for the study of the claim process in the presence of covariates the generalized linear models. Based on the concepts of quasi-likelihood and extended quasi-likelihood, Renshaw pays attention to the variety of probability distributions that are available and to the parameter estimation and model fitting techniques that can be used for the claim frequency and the claim severity process.(For further information on this, confer to Frangos and Vrontos (2001)).

2.3.4 Properties of the Generalized BMS

- It is fair .It takes into account the number of claims, the claim severity and the significant a priori rating variables for the the number of claims and for the claim severity for each policyholder.
- It is financially balanced for the insurer. Each year the average premium will be equal to

$$P = \exp(c_i^{t+1}\beta^{t+1})\exp(d_i^{t+1}\gamma^{t+1}) \qquad (7)$$

- All the properties for generalized BMS without the a priori rating variables hold for this BMS as well. Policy-holders with the same characteristics are paying the same premium equal to (7) in the beginning.

- Premium increases proportionally to the number and to the severity of the claims.
- Premium always decreases when no accidents occur.
- Phenomenon of bonus hunger will decrease by this generalized BMS
- The introduction of the severity component is more crucial than the number of claims for the insurer.
- Premiums vary with the variables that affect the distributions of claims number and severity.

.

CHAPTER 3

DESIGN OF AN OPTIMAL BONUS-MALUS SYSTEM WHERE THE CLAIM FREQUENCY DISTRIBUTION IS GEOMETRIC AND THE CLAIM SEVERITY DISTRIBUTION IS PARETO

3.1 Introduction

In chapter 3, under the assumptions that the claim number follows a Geometric distribution and that the size of the claims follows a Pareto distribution, we will design both the classical optimal bonus-malus system, under which the premium that is assigned to each policyholder is based only on his claims number neglecting their size, and an optimal bonus-malus system, under which the premium is set by taking into account both the frequency and the severity of the claims.

3.2 Design of optimal BMS based on the frequency component

Mixed Poisson distributions have thicker tails than the Poisson distribution, thus they provide a better fit to claim frequency data than the Poisson distribution when the portfolio is heterogeneous i.e. all policyholders will have a constant but unequal underlying risk of having an accident. That is, the expected number of claims differs from policyholder to policyholder. In chapter2, we used a Negative Binomial distribution for modeling claim frequency, for the same job, in this chapter we will use the Geometric distribution - which is a mixture of a Poisson and an Exponential distributions.

For a heterogeneous portfolio, it is assumed that the number of claims *k*, given the parameter λ, is considered to follow the Poisson (λ) distribution:

$$p_{\lambda}(k \mid \lambda) = \frac{\lambda^{k}}{k!} e^{-\lambda} \tag{1}$$

k =0,1,2,3,… and $\lambda>0$ where λ denotes the different underlying risk of each policyholder to have an accident. Also it is assumed that λ~Exponential (θ), so:

$$u(\lambda) = \theta e^{-\theta\lambda}, \lambda > 0 \tag{2}$$

From (1), (2) it can be proved that the unconditional distribution of the number of claims k will be a Geometric distribution:

$$P(k) = \int_0^\infty \mathrm{P}_\lambda(k \mid \lambda) u(\lambda) d\lambda = \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!} \theta e^{-\lambda\theta} d\lambda = \ldots = \theta(1+\theta)^{\kappa+1}, k = 0,1,2,\ldots \tag{3}$$

If $K=\sum_{i=1}^{t} k_i$ is the total number of claims that a policyholder had in t years, with k_i the number of claims that the policyholder had in the year i, i = 1,...,t. Then, for a given λ, the conditional distribution of K is :

$$P(k_1, k_2, \ldots, k_t \mid \lambda) = \frac{e^{\lambda t}\lambda^K}{\prod_{i=1}^{t} k_i!} \tag{4}$$

From (3), (4) and by applying the Bayes' theorem, the posterior structure function of λ for a policyholder or a group of policyholders with claim history k_1,....k_t, is obtained in the following way:

$$U(\lambda \mid k_1, k_2, \ldots, k_t) \propto P(k_1, k_2, \ldots, k_t \mid \lambda) U(\lambda) \propto e^{-\lambda t}\lambda^K e^{-\lambda\theta} \propto e^{-\lambda(t+\theta)}\lambda^K$$

If, for a constant, $\int_0^\infty A e^{-\lambda(t+\theta)}\lambda^K d\lambda = 1$ then $A = \frac{(t+\theta)^{\mathrm{K}+1}}{\Gamma(\mathrm{K}+1)}$

$$U(\lambda \mid k_1, k_2, \ldots, k_t) = \frac{(t+\theta)^{\mathrm{K}+1}}{\Gamma(\mathrm{K}+1)} e^{-\lambda(t+\theta)}\lambda^K, \lambda > 0 \tag{5}$$

The optimal choice of λ_{t+1} for a policyholder with claim history k_1, ...k_t is obtained by the use of the quadratic error loss function and it will be the mean of the posterior structure function, that is:

$$\hat{\lambda}_{t+1} = \frac{\mathrm{K}+1}{t+\theta} = \bar{\lambda}\frac{\mathrm{K}+1}{t\bar{\lambda}+1}, \bar{\lambda} = \frac{1}{\theta} \tag{6}$$

3.2.1 Calculation of the Premium according the Net Premium Principle

If, only the claims number is taken into account to set the premium, then assuming that at time t=0 the initial premium is 100, the premium at time t+1 is:

$$\mathrm{Pr}\,emium = 100\frac{\mathrm{K}+1}{t\bar{\lambda}+1} \tag{7}$$

3.3 Design of optimal BMS with a frequency and a severity component based on the a posteriori criteria

For modeling claim frequency, we proposed that the number of claims is distributed according to the Geometric distribution For modeling claim severity we will use the Pareto Distribution (Exponential-Gamma mixture) like we did in chapter 2.Actually, in many actuarial papers long tail distributions are used to model claim severity because in addition to many small claim severities, high claim severities can also be observed.

Next, let us consider a quick review of the Pareto distribution for modeling claim amount. Assume that the amount of the claim of each insured x is distributed according to the Exponential distribution with a given parameter y- the mean claim size for each insured and assume that y ~ IG(m,s) as y is not the same for all policyholders Then $x|y$ is the exponential distribution with parameter y. Thus it can be prove that if $x|y$ ~ *Exponential*(y) and y ~ Inverse *Gamma*(s,m) then x ~ *Pareto*(s,m).Then, by applying Bayes' theorem we obtain the posterior distribution of the mean claim size y given the claims size history of the policyholder x_1, ..., x_K and it is the I G($s + K$, m+$\sum_{k=1}^{K} x_k$).Using the quadratic loss function we obtain:

$$\hat{y}_{t+1} = \frac{m + \sum_{k=1}^{K} x_k}{s + K - 1} \tag{8}$$

,optimal choice of y_{t+1} the mean of the posterior distribution of the mean claim size.

3.3.1 Calculation of the Premium according the Net Premium Principle

If the claims number and the claims amount is taken into account to set the premium, then, for a policyholder with claim frequency history $k_1, k_2, ..., k_t$ and claim severity history x_1, ..., x_K , from (6) and (8) ,using the net premium principle the premium at time t+1 is:

$$\Pr emium = \frac{K+1}{t+\theta} \cdot \frac{m + \sum_{k=1}^{K} x_k}{s + K - 1}.$$

Finally, as in chapter 2, it should be mentioned that optimal BMS based on the claim frequency and claim severity component differentiates the premiums according to the claim severity and thus it is fairer for the policyholders.

CHAPTER 4:

USING MIXED POISSON PROCESSES IN CONNECTION WITH BONUS-MALUS SYSTEMS

4.1 Introduction

In this chapter, we propose a parametric method encompassing those encountered in the literature for the construction of BMS. Furthermore, this parametric method is compared with a non-parametric one that permits a simple formulation of the stationary and transition probabilities in a portfolio for the construction of a BMS with finite number of classes

The number of car accidents is Poisson distributed for each risk in the portfolio which is heterogeneous, the frequency of the risks differs from each other and it is assumed to follow a random variable.

- *N(t)* is the number of claims in (0,t]
- $P(k,t) = P_k(t) = P[N(t) = k]$ is the probability that a risk causes *k* accidents in *t* years we have

$$P(k,t \mid \Lambda) = e^{-\Lambda t} \frac{(\Lambda t)^k}{k!} \tag{1}$$

$$P_k(t) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^k}{k!} dU(\lambda)$$

(2)

,Λ is a r.v. with cumulative density function *U(λ)*.

Lemaire (1985) used a Gamma distribution for Λ, so *N(t)* followed a Negative Binomial distribution and the construction of a BMS was easy. Tremblay (1992) used the Inverse Gaussian distribution for Λ thus *N(t)* followed a Poisson Inverse Gaussian distribution. In this setting, the construction of a BMS is complicated needing for example the use of modified Bessel functions which is in fact unnecessary.

Here, as we have already mentioned, we will use a more general parametric distribution and a non-parametric distribution to fit an automobile portfolio and we will construct a BMS using the Bayes theorem and the form of the mixed Poisson distribution.

- Properties that will be used:

$$\mathrm{M}_{N(t)}(s) = P(0, t - te^{s}) \tag{3}$$

$$\mathrm{M}_{\Lambda}(s) = P(0, -s) \tag{4}$$

, $\mathrm{M}_X = \mathrm{E}[e^{s\mathrm{X}}]$ is the moment generating function of *X*.

The claim number *N(t)* is a pure birth process, so the estimation of the intensity of the process

$$E[N(t+1) - N(t) \mid N(t)]$$

needs data reported on a long period of time. This is not always the case in practice and the model designed by Walhin and Paris (1999) ,allows the estimation of the intensity and the construction of a tarification based only on the total number of claims . This system is comparable to a BMS with an infinite number of classes. Also, because the process is stationary:

$$EE[N(t+1) - N(t) \mid N(t)] = EN(1)$$

i.e. the system is at the equilibrium which is not the case with the BMS we come across in practice.

4.2 Parametric estimation

A three parameters distribution is used, encompassing the Negative Binomial and the Poisson Inverse Gaussian distributions. This is Hofmann's distribution (Hofmann

(1955)), that has also been discussed by Kestemont and Paris (1985). Hofmann's distribution is based on a model that is verified on the data that Walhin and Paris (1999) are studying i.e. Poisson distribution rejected, infinite divisibility, Shaked theorem thus it leads to an excellent fit.

Hofmann's distribution is defined as follows:

$$P(0,t) = e^{-\theta(t)}$$

$$P(\kappa,t) = (-1)^k \frac{t^k}{k!} P^{(k)}(0,t) \text{ , k=1,2,...}$$

$$\theta'(t) = \frac{p}{(1+ct)^a} \text{ , p>0, c>0, } \alpha \geq 0$$

$$\theta(0) = 0$$

By integration,

$$\theta(t) = pt \quad \text{if } \alpha\text{=0}$$

$$\theta(t) = \frac{p}{c}\ln(1+ct) \quad \text{if } \alpha\text{=1}$$

$$\theta(t) = \frac{p}{c(1-a)}\left[(1+ct)^{1-a} - 1\right] \text{ else}$$

the following distributions are encompassed: Poisson $(a = 0)$, the Negative Binomial $(a = 1)$ and the Poisson Inverse Gaussian $(a = \frac{1}{2})$.

Panjer's algorithm (Panjer (1981)) is used to write a recursion for the probabilities of *N{t)*. That is because *N(t)* can be interpreted as a Compound Poisson distribution due to the use of infinite divisibility arguments (for more details refer to Maceda (1948), Feller (1971), Kestemont and Pans (1985).

$$P(0,t) = e^{-\theta(t)}$$

$$(k+1)P(\kappa+1,t) = \frac{pt}{(1+ct)^a}\sum_{i=0}^{k}\frac{\Gamma(a+i)}{\Gamma(a)i^t}\left(\frac{ct}{1+ct}\right)^i P(\kappa-i,t) \tag{5}$$

N(t) has mean and variance:

$$E(N(t)) = pt$$
$$V(N(\mathrm{T})) = pt + pcat^2$$

The loglikelihood to be maximized is:

$$l(p,c,a) = \sum_{i=0}^{\infty} N_i \ln\{P(i,1)\}$$

, for N_i the number of policies with i claims.

$\hat{\mathrm{E}}(N(1)) = \hat{p} = \overline{N}$, where $\overline{N}$ is the experimental mean (for more see Hurlimann (1990)).The other two parameters $\hat{c}, \hat{a}$ are estimated by a numerical maximization.

4.3 Non-parametric estimation

Here, a mixed Poisson distribution is used for which we don't specify a parametric distribution *U(λ)* for Λ,unlike the parametric case that a *u* function has been chosen which remained to estimate the parameters.

We will attain the maximum likelihood estimate of *U* for a discrete distribution function *U(λ)* with a maximum number *m* of growing points (for more see Simar (1976)).

The probabilities $P(k,t)$ are given by:

$$P(k,t)=\sum_{j=1}^{m} p_j e^{-\lambda_j t}\frac{(\lambda_j t)^k}{k!} \tag{6}$$

and $\sum_{j=1}^{m} p_j = 1$, $p_j \geq 0 \;\; \forall j$ and m , the number of support points, i.e. the number of homogeneous classes of risks.

In the sequel it is supposed that: $0 \leq \lambda_1 < \lambda_2 < ... < \lambda_m$.

An algorithm for the calculation of the non-parametric maximum likelihood estimators for an automobile portfolio is provided by Simar (1976). Nevertheless, the algorithm does not converge certainly to the global maximum because the loglikelihood is not concave everywhere and Simar (1976), in his numerical example, did not verify the valid property for all non-parametric mixtures of the exponential family (see Lindsay (1995)).

If N is the maximum number of claims per risk:

$$\hat{\mathrm{E}}(N)=\overline{N} \tag{7}$$

We can also find the maximum likelihood using a Newton-Raphson technique and simplify the procedure by reducing the number of parameters to be estimated due to property (7).

For the number of mass points *m,* it has been shown by Simar (1976) that, the maximum likelihood estimator can be unique under the following conditions:

$$m \leq \min\left(q,\left[\frac{N+2}{2}\right]\right) \text{ if } \lambda_1=0$$

$$m \le \min\left(q, \left[\frac{N+1}{2}\right]\right) \text{ if } \quad \lambda_1 > 0$$

, q is the number of classes for which the observation is different from 0.

At this point, we should mention that the maximum likelihood method we have described gives more information than the good guy bad guy model of Lemaire (1985). That is because it gives us the number of mass points needed to have the highest likelihood while Lemaire's model imposes two mass points. Also, the non-parametric case gives a physical interpretation of the heterogeneity of the portfolio

Finally, in Walhin's and Paris's portfolio the fit is excellent (Chi-Square statistic is 0.1245 with 6 classes and the fit is excellent because there are 5 free parameters for 7 classes Note that the distribution is not infinitely divisible in this case. For more see Walhin and Paris(1999)).

4.4 Optimal bonus-malus systems

Bonus- malus systems depend on the history of the claims $k_1, k_2, ..., k_t$ caused by the policyholder, in this model it is sufficient to consider the total number of claims $K=\sum_{i=1}^{t} k_i$ without reference to their history:

$$du(\lambda \mid N(t)-N(t-1)=k_t, ..., N(1)-N(0)=k_1)$$
$$= \frac{P(N(t)-N(t-1)=k_t, ..., N(1)-N(0)=k_1 \mid \lambda) dU(\lambda)}{P(N(t)-N(t-1)=k_t, ..., N(1)-N(0)=k_1)}$$
$$= \frac{\dfrac{e^{-\lambda t}\lambda^h}{\Pi_{i=1}^{t} k_i} dU(\lambda)}{\displaystyle\int_0^\infty \frac{e^{-\lambda t}\lambda^h}{\Pi_{i=1}^{t} k_i} dU(\lambda)}$$

For the first year, the premium is an a priori premium because there is no information concerning the risk:

E(N(1)) = E(Λ)

For the year t the information which consists in the number of accidents during the first t years is taken into account, because the history of the accidents is unimportant, and the premium is:

$$E[N(t+1)-N(t) \mid N(t)=K] = E(\Lambda \mid N(t)=K)$$
$$=\frac{K+1}{t}\frac{P(K+1,t)}{P(K,t)}$$

In the binomial negative case, this expression can be reduced and it is simpler to use than the formulae derived by Tremblay (1992) for the particular case of the Poisson Inverse Gaussian distribution more specifically it can be written as:

$$\frac{p+Kc}{1+ct} = p\frac{1}{1+ct}+\frac{K}{t}\frac{ct}{1+ct}$$

Depending on K and t, a bonus-malus table can be constructed assuming that the first premium paid is 100, with:

$$\frac{100}{E(\Lambda)}\frac{K+1}{t}\frac{P(K+1,t)}{P(K,t)} \qquad (8)$$

4.4.1 Properties of the Optimal BMS

- The system is financially balanced. Every year:

$$\sum_{K=0}^{\infty} P(K,t) E(\Lambda \mid N(t)=K) = \mathrm{E}(\Lambda) \quad \forall t$$

- Premium always increases when more accidents occur
  $E(\Lambda \mid N(t)=K+1) > E(\Lambda \mid N(t)=K) \quad \forall t,k$ (proof by Cauchy-Schwartz inequality)

- Premium decreases when no more accidents occur:

$$\frac{d}{dt} E(\Lambda \mid N(t)=K) \leq 0 \quad \forall t,k \qquad \text{(proof by Cauchy-Schwartz inequality)}$$

<u>4.5 Bonus- malus system for loaded premiums</u>

With the principle of zero utility we can construct a BMS for charged premiums, as Lemaire (1985) and Tremblay (1992), using an exponential utility function:

$$u(x) = \frac{1}{\gamma}(1 - e^{-\gamma x}) \quad ,\gamma > 0$$

The a priori premium by the use of (3) and (4) can be written (see Gerber (1979)):

$$P = \frac{1}{\gamma} \ln E[e^{\gamma N(1)}]$$

$$= \frac{1}{\gamma} \ln E[e^{w\Lambda}] \ , \ w = e^{\gamma} - 1$$

The a posteriori premium can be calculated like in the previous section:

$$
\begin{aligned}
P &= \frac{1}{\gamma}\ln E[e^{w\Lambda} \mid N(t)=K] \\
&= \frac{1}{\gamma}\ln\left\{\frac{1}{P(K,t)}\int_0^\infty e^{w\Lambda} e^{-\lambda t}\frac{(\lambda t)^K}{K!}dU(\lambda)\right\} \\
&= \frac{1}{\gamma}\ln\left\{\left(\frac{t}{t-w}\right)^K \frac{P(K,t-w)}{P(K,t)}\right\}
\end{aligned}
$$

A bonus-malus table can be constructed, by normalizing assuming that the first premium paid is 100, with:

$$
100\frac{\ln\left\{\left(\frac{t}{t-w}\right)^K \frac{P(K,t-w)}{P(K,w)}\right\}}{\ln\left\{P(0,-w)\right\}} \text{, for } w = e^{\gamma}-1 \qquad (9)
$$

4.6 Notes

4.6.1 Parametric Case

Walhin and Paris used the expected value principle to obtain the optimal BMS table with the Hofmann fit and it is comparable with Lemaire's (1985) and Tremblay's (1992) (for more see Walhin and Paris (1999)).
Due to the form we chose for $\theta'(t)$ we will always have:

$$\lim_{t\to\infty} E(\Lambda \mid N(t)=0)=0$$

and

$$\lim_{t\to\infty} E(\Lambda \mid N(t)=K)=0 \quad \forall K$$

This is related to the observation that when the frequency is low most of the drivers are in the cheapest class .Also, because the convergence to 0 is attained by far after the mean driving time, in practice there is no way for a driver to pay a premium equal to zero. Nevertheless, the problem could be solved by adding a constant premium in our model:

$$\theta'(t)=\delta+\frac{p}{(1+ct)^{a}} \tag{10}$$

From (10), the driver always pays a minimum premium δ because the premium has the following asymptotic behavior:

$$\lim_{t\to\infty} E(\Lambda \mid N(t)=K)=\delta \quad \forall K$$

We should also mention that for typical automobile portfolios the following property seems valid:

$$\frac{d^2}{dt^2}E(\Lambda \mid N(t)=K)\geq 0 \quad \forall K$$

Finally, the optimal BMS tables for loaded premiums, obtained with the zero-utility principle by Walhin and Paris, are also comparable with Lemaire's (1985) and Trcmblay's (1992) even for unreasonable values of γ (for more see Walhin and Paris (1999)).

4.6.2 Non- Parametric Case

The formulae (8) and (9) can be rewritten as:

$$(8) = \frac{100}{E(\Lambda)} \frac{\sum_{j=1}^{m} p_j e^{-\lambda_j t} \lambda_j^{K+1}}{\sum_{j=1}^{m} p_j e^{-\lambda_j t} \lambda_j^{K}} \tag{11}$$

$$(9) = \frac{100}{\ln\{P(0,-w)\}} \frac{\sum_{j=1}^{m} p_j e^{-\lambda_j (t-w)} \lambda_j^{K}}{\sum_{j=1}^{m} p_j e^{-\lambda_j t} \lambda_j^{K}} \tag{12}$$

The asymptotic behavior of the BMS tables constructed by J.F. Walhin and J.Paris with the non- parametric fit, can be described as follows:

If $\lambda_1 > 0$

$$\lim_{t\to\infty}(11) = \min_{\lambda_j} \lambda_j \frac{100}{E(\Lambda)}$$

$$\lim_{t\to\infty}(12) = \min_{\lambda_j} \lambda_j \frac{w}{\ln P(0,-w)}$$

If $\lambda_1 = 0$

$$\lim_{t\to\infty}(11) = \min_{\lambda_j > 0} \lambda_j \frac{100}{E(\Lambda)} \quad \text{if } K > 0$$

$$= 0 \qquad \text{if } K = 0$$

$$\lim_{t\to\infty}(12) = \min_{\lambda_j > 0} \lambda_j \frac{w}{\ln P(0,-w)} \qquad \text{if } K > 0$$

$$= 0 \qquad \text{if } K = 0$$

We must point out that there is a great difference between the BMS tables constructed with the non- parametric fit and the parametric case tables. More specifically, the discontinuity of Λ is reflected and that gives the BMS table a curious comportment with local almost constant premiums. The same comments hold for the charged premium BMS table. We have these results for the non-parametric method because the estimation of the distribution function u of the random variable Λ is based only on the observation of *N(l)*. Even if the observation has a longer period, the trouble remains and that is because, due to the fact that the frequency is low, the number of points on increase of u is always low and the number of classes of risks is low. (for more see Walhin and Paris (1999))

So, for the construction of a BMS table, the parametric approach should be preferred because of its 'continuity. However, the non-parametric fit is interesting for the evaluation of the mean of difficult functions of Λ over the portfolio because complicated numerical integration is replaced by a short summation.

4.7 The stationary and transient distributions of the policyholders in a BMS with finite number of classes

As we have seen, the non-parametric model was too discontinuous to give a nice form for the premiums. Thus, a parametric Mixed Poisson model should be preferred for the construction of optimal BMS. The non-parametric fit for Mixed Poisson distributions presents high interest for the evaluation of the mean of difficult functions over the risk's portfolio.

Assume that *f(λ)*, is a complicated function of λ. If we want to calculate

$$E(f(\lambda)) = \int_0^\infty f(\lambda)dU(\lambda) \tag{13}$$

even a numerical integration can be unworkable.

Nevertheless, the use of the non-parametric structure function of Λ will be more efficient. It will give a better fit and no numerical integration will be needed. A convex combination of some $f(\lambda)$ will be performed and the relation (13) will become:

$$E(f(\lambda)) = \sum_{j=1}^{m} f(\lambda_j)p_j$$

(14)

, most of the time (14) will be a sum with 3 of 4 terms.

This technique can be applied to find the stationary and transient distributions of a typical Markov chain used in automobile insurance.

As we know, the BMS with finite numbers of classes are characterized by *s*+1 classes with growing premium percentages b_i, *i=0,…,s.* The transition rules depending on the number of accidents caused during one year, give the movements of the drivers between the classes. Also, the transition rules give the model the Markov property almost every time. Even when it is not the case, a redefinition of the classes can give the Markov property to the model (see Lemaire (1985) for an application to the BMS). If we define as X_t the class at time t and as Y_t the transition rules then: $X_t = X_{t-1} + Y_t$.

Assuming that Y_t are independent random variables, we can find the stationary distribution of the risks in the BMS using Dufresne's (1988) recursive technique. At

this point, we should mention that for BMS with nonuniform penalties per claim Dufresne's technique would not be applicable while the technique described hereunder remains applicable for every BMS.
The general Mixed Poisson process can't be used for *N(t)* due to the independence condition. However a Poisson process is adequate. For the calculation of the stationary distribution of risks with mean λ, Dufresne (1995) used the Poisson distribution and found it recursively, as a function of λ. Due to the fact that he finds complicated functions of λ, a software handling symbolic computations is welcome.
Then, denoting as $F^{\lambda}(x,s)$ the stationary distribution with x as the class, Dufresne (1995) proposed to find the unconditional stationary distribution by:

$$F^{\lambda}(x,s) = \int_{0}^{\infty} F^{\lambda}(x,s)\,dU(\lambda)$$

If we have a non-parametric fit for *N(t)*, this integration will be very easy as we have already mentioned. On the second hand, if only a few values of λ are needed, instead of using the algorithm of Dufresne (1988) we can use the traditional technique of norming the left eigenvector of the transition probability matrix.

Regardless the method we will use, we can calculate the stationary probability vector for the values of λ and their weighted mean which is the stationary probability vector of the portfolio .Also, we have to mention that all this remains true for the transient distribution functions. Note that the problem encountered by Coene and Doray (1996) ,who used simulation to find the stationary distribution of a portfolio with *N(t)* negative binomial distributed, is solved by this technique. Finally by taking powers of the transition probability matrix, we also find transient probabilities (for more see the application of Walhin and Paris (1999)).

CHAPTER 5:

GENERAL PROPERTIES OF MIXED POISSON DISTRIBUTIONS APPLICATION IN THE CASE OF AUTOMOBILE COLLISION CLAIMS DISTRIBUTIONS

5.1 Introduction

As we know, for the construction of optimal BMS in automobile insurance, the distribution of the number of car accidents is frequently chosen within the "mixed-Poisson" family. In this chapter, we will show the general properties of "mixed Poisson" family distributions and we will give a unifying approach of several particular cases including the geometrical, the P-Erlang, the Negative Binomial and the P-inverse Gaussian distributions. The problem of adjustment remains the thickness of the tails of the underlying distributions. In order to avoid this problem we will introduce a new family of "mixed-Poisson", built upon "fatty-tailed" underlying distributions, the so called "P-rational'' distributions.

5.2 Main results of the mixed Poisson distributions

We will denote as X the number of claims occurring in a unit period and as in the preceding, Λ will be the common risk parameter giving the expected number of claims during the same unit period. In a given heterogeneous portfolio, Λ is a random variable with cumulative density function *$U(\lambda)$.*

For a given risk λ, the random variables $P_\lambda(k|\lambda)$ *($k = 0,1,\ldots$)* which give the number *(k)* of claims for one period, are identically distributed and follow a Poisson distribution.

The probability that *k* claims are observed in one period is then given by:

$$P(X = k) \equiv p(k) = \int_0^\infty p_\lambda(k \mid \lambda) * dU(\lambda)$$

with

$$p_{\lambda}(k \mid \lambda) = \frac{\lambda^{k}}{k!} e^{-\lambda}$$

5.2.1Moments of mixed distributions

- The mean of a mixed distribution is equal to the mean of the underlying distribution$U(\lambda)$.

$$E[X] = E[\lambda]$$

- $Var[X] = E(X^2) - [E[X]]^2 = E(\Lambda) + Var(\Lambda)$

(for the proof of both lemma's see Hulin and Justens (Banque Pictet-Luxembourg HEFF/Cooremans Bruxelles)) .

5.2.2 Negative Binomial distribution

Transition Probabilities

We assume the underlying distribution follows a gamma distribution (this model is that of Lemaire (1977)).

$$dU(\lambda) = u(\lambda) = \frac{\exp(-\tau\lambda)\lambda^{\alpha-1}\tau^{\alpha} d\lambda}{\Gamma(\alpha)}$$

We obtain:

$$P(k)=\frac{\tau^{a}}{\Gamma(\alpha)k!}\int_{0}^{\infty}e^{-(\tau+1)\lambda}\lambda^{k+\alpha-1}d\lambda$$

For $\lambda^{'}=\lambda(\tau+1)$ we obtain:

$$P(k)=\frac{\tau^{a}}{\Gamma(\alpha)k!(\tau+1)^{a+k}}\int_{0}^{\infty}e^{-\lambda^{'}}\lambda^{'(k+\alpha-1)}d\lambda^{'}$$

and the mixed distribution is a Negative Binomial (α,τ)

$$P(k)=\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)k!}\frac{\tau^{a}}{(\tau+1)^{k+a}}$$

<u>Moments</u>

It can be proved that:

$$E(k)=E(\Lambda)=\frac{1}{\Gamma(\alpha)\tau}\cdot\alpha\cdot\Gamma(\alpha)=\frac{\alpha}{\tau}=E(X)$$

<u>Proof:</u> $E(\Lambda)=\int_{0}^{\infty}\lambda f(\lambda)d\lambda=\frac{\tau^{\alpha}}{\Gamma(\alpha)}\int_{0}^{\infty}\lambda^{\alpha}e^{-\lambda\tau}d\lambda$, for $\lambda'=\tau\lambda$ and integrating by parts

$$E(\Lambda)=\frac{1}{\Gamma(a)\tau}\int_{0}^{\infty}\lambda'^{a}e^{-\lambda'}d\lambda'=[-e^{-\lambda'}\lambda'^{a}]_{0}^{\infty}+a\int_{0}^{\infty}\lambda'^{(a-1)}e^{-\lambda'}d\lambda'$$

For the variance,

$$V(k)=Var[X]=E(X^{2})-[E[X]]^{2}=E(\Lambda^{2})+E(\Lambda)-[E(\Lambda)]^{2}=\frac{a(\tau+1)}{\tau^{2}}$$

Proof: We have to calculate $E(\Lambda^2) = \int_0^\infty \lambda^2 f(\lambda) d(\lambda) = \frac{\tau^a}{\Gamma(a)} \int_0^\infty \lambda^{a+1} e^{-\lambda\tau} d\lambda$

for $\lambda' = \tau\lambda \; E(\Lambda^2) = \frac{(a+1) * \Gamma(a+1)}{\Gamma(a)\tau^2}$

and as $\Gamma(\alpha+1)=\alpha!$ we come to $E(\Lambda^2) = \frac{a(a+1)}{\tau^2}$

5.2.3 Geometrical distribution

Transition Probabilities

We assume the underlying distribution follows a negative exponential distribution (for more see Hulin (1999) and Justens (1996)).

$$P(k) = \int_0^\infty \mathrm{P}_\lambda(\kappa \mid \lambda) u(\lambda) d\lambda = \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!} \theta e^{-\lambda\theta} d\lambda = \ldots = \theta(1+\theta)^{\kappa+1}, \kappa = 0,1,2,\ldots$$

Proof: Using the lemma $\int x^k e^{-x(\theta+1)} dx = e^{-x(\theta+1)} \left[\sum_{j=0}^{k} \frac{\frac{k!}{(k-j)!} x^{k-j}}{(\theta+1)^{j+1}} \right]$

we obtain $P(k) = \frac{\theta}{k!} \left[-e^{\lambda(\theta+1)} \cdot \sum_{j=0}^{k} \frac{\frac{k!}{(k-j)!} \lambda^{k-j}}{(\theta+1)^{j+1}} \right]_0^\infty = \frac{\theta}{(\theta+1)^{k+1}}$ (for more see Hulin and Justens (Banque Pictet-Luxembourg HEFF/Cooremans Bruxelles)).

Moments

It can be proved that:

$$E(k) = E(\Lambda) = \frac{1}{\theta} = E[X]$$

Proof: $E(\Lambda) = \int_0^\infty \lambda^2 f(\lambda) d\lambda = \int_0^\infty \theta \cdot \lambda \cdot e^{-\theta\lambda} d\lambda$

For the variance,

$$V(k) = Var(X) = \frac{1+\theta}{\theta^2}$$

Proof: We have $E(\Lambda^2) = \int_0^\infty \lambda^2 f(\lambda) d\lambda = \int_0^\infty \lambda^2 \cdot \theta \cdot \lambda \cdot e^{-\theta\lambda} d\lambda = \frac{2}{\theta^2}$

5.2.4 Poisson-Erlang distribution

Transition Probabilities

We assume the underlying distribution follows an Erlang distribution with $U(\lambda) = \theta^2 \cdot \lambda \cdot e^{-\theta\lambda}$ (for more see Hullin (1999) and Justens (1996)).

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

Proof: Using again the lemma $\int x^k e^{-x(\theta+1)} dx = e^{-x(\theta+1)} \left[\sum_{j=0}^{k} \frac{\frac{k!}{(k-j)!} x^{k-j}}{(\theta+1)^{j+1}} \right]$

we obtain $P(k) = \dfrac{\theta^2}{k!}\left[-e^{\lambda(\theta+1)} \cdot \sum_{j=0}^{k+1} \dfrac{\dfrac{k!}{(k-j)!}\lambda^{k-j}}{(\theta+1)^{j+1}} \right]_0^{\infty} = \dfrac{\theta^2 (k+1)}{(\theta+1)^{k+2}}$.

Moments

It can be proved that

$$E(k) = E(\Lambda) = \frac{2}{\theta} = E[X]$$

Proof: $E(\Lambda) = \int_0^{\infty} \lambda^2 f(\lambda) d\lambda = \int_0^{\infty} \theta^2 \cdot \lambda^2 \cdot e^{-\theta\lambda} d\lambda$

For the variance,

$$V(k) = Var(X) = \frac{2(1+\theta)}{\theta^2}$$

Proof: We have $E(\Lambda^2) = \int_0^{\infty} \lambda^2 f(\lambda) d\lambda = \int_0^{\infty} \lambda^2 \cdot \theta^2 \cdot \lambda \cdot e^{-\theta\lambda} d\lambda = \dfrac{6}{\theta^2}$

5.2.5 Poisson-Inverse Gaussian distribution

Transition Probabilities

Finally we consider the Poisson Inverse Gaussian Distribution (for more see Besson (1992) and Trembley (1992))

We have:

$$dU(\lambda) = \frac{\mu}{\sqrt{2\pi\beta\lambda^3}} \cdot e^{-\left(\frac{(\lambda-\mu)^2}{2\beta\lambda}\right)} d\lambda$$

As there is no general formula:

For $k = 0$ we have:

$$P(0) = \int_0^\infty e^{-\lambda} * \frac{\lambda^k}{\kappa!} \frac{\mu}{\sqrt{2\pi\beta\lambda^3}} e^{-\left(\frac{\left((\lambda-\mu)^2\right)}{2\beta\lambda}\right)} d\lambda = e^{\frac{\mu}{\beta}\left[1-\sqrt{1+2\beta(1-\kappa)}\right]}$$

For $k = 1$ we have:

$$P(1) = \mu P(0)(1+2\beta)^{-\frac{1}{2}}$$

For $k > 1$ we have:

$$(1+2\beta)\kappa(\kappa-1)P(k) = \beta(\kappa-1)(2\kappa-3)P(k-1) + \mu_{\kappa-2}^2 P(k-2)\ .$$

<u>Moments</u>

We have

$$E(k) = E(X) = \mu$$

$$V(k) = Var(X) = \mu(1+\beta)$$

(General results are given by Besson (1992) and Trembley (1992))

<u>5.3 Geometrical distributions with two parameters</u>

We generalize the geometrical distribution (for more see LEMAIRE, J. - ZI HONGMIN (1994)) by means of conditional probabilities. Let:

$$P(0) = 1 - \frac{1}{1+a} = 1 - a$$

It can be written:

$$P(k) = a^k (1-a)$$

a is the probability to observe at least one claim during the period. We calculate:

$$P_{1claima/atleast1} = 1 - a$$

So the probability to observe a second claim is also equal to a. But this cannot be the case and these probabilities cannot be equal because, as the first claim occurs during the time period, there is less time left for the second to occur. We will assume that:

$$P_{1claima/atleast1} = 1 - \beta$$

We construct the probability distribution:

$$P(0) = 1 - a \quad P(k) = \alpha \beta^{(\kappa-1)} (1-\beta) \quad k \in N_0^+$$

with mean and variance

$$\mu = \frac{a}{1-\beta} \text{ and } \sigma^2 = \frac{\alpha(1+\beta-\alpha)}{(1-\beta)^2} .$$

Finally, from the equality of the first two moments, the distribution is easily adjusted:

$$\alpha = \frac{2\overline{x}^2}{\overline{x}^2 + \overline{x} + s^2} \quad \beta = \frac{\overline{x}^2 - \overline{x} + s^2}{\overline{x}^2 + \overline{x} + s^2} .$$

The basic problem that we have is the thickness of the tails of the underlying distributions. The systematic use of a negative exponential isn’t sufficient enough for a great number of claims to occur .In order to avoid this problem, we will introduce the “P-rational distributions, a new family of “mixed-Poisson”, built upon “fatty-tailed” underlying distributions.

5.4 Rational distributions

5.4.1 General properties

Consider a cumulative distribution defined on $\mathbb{R}^+$ with

$$U(\lambda) = \frac{g(\lambda)}{g(\lambda) + a}$$

, where *g(λ)* is a strictly increasing function of λ on $\mathbb{R}^+$ which goes to infinity when λ goes to infinity, and α is a positive constant. We may also assume that *g(λ)* is strictly decreasing, going to minus infinity, and that a is strictly negative. In both cases, it can be easily proved that *U(λ)* is a cumulative function. We also have:

$$U(\lambda) = 1 - \frac{a}{g(\lambda) + a}$$

and it shows that with respect to $\frac{1}{g(\lambda)}$ the distribution tail tend to 1. We just have to select the desired thickness of the curve tail.

Also we notice that the density function can be written as:

$$u(\lambda) = \frac{a \cdot g'(\lambda)}{\left(g(\lambda) + a\right)^2}$$

If we have a polynomial form for *g(λ),* it has to be at least a quadratic one as we obviously want to get a distribution with a finite first order moment. The first derivative must be at least of the first order so that:

$$u(0) = 0$$

This property shows that there are no drivers with a totally riskless behavior.

At this point, we should mention that in the data fitting this kind of function never leads to a handsome form for the claim distribution but numerical values can be easily given with mathematical packages. Finally, the variance is not necessarily needed in data fitting because distributions with infinite theoretical variance sometimes better fit observed data.

5.5 Particular cases

5.5.1 Quadratic case

We must take the most simple form of quadratic function for *g(λ)* in order to get at least a first order moment which permits a data fitting. We assume:

$$g(\lambda) = \lambda^2$$

we get:

$$U(\lambda) = \frac{\lambda^2}{\lambda^2 + \alpha}$$

$$u(\lambda) = \frac{2\alpha\lambda}{(\lambda^2 + \alpha)^2}$$

The comportment of the model and the general form of the distribution are given by the second order derivative and we can see the evolution of the mode and the thickness of the tail as an increasing function of a.

$$u''(\lambda) = \frac{2\alpha(\alpha - 3\lambda^2)}{(\lambda^2 + \alpha)^3}$$

For the first order moment, we have:

$$\int \frac{\lambda^2}{\lambda^2 + a} d\lambda = -\frac{\lambda}{2 \cdot \left(\lambda^2 + \alpha\right)} + \frac{\arctan \frac{\lambda}{\sqrt{a}}}{2\sqrt{a}}$$

so we get:

$$E[\Lambda] = \int_0^\infty \lambda f(\lambda) d\lambda = \int_0^\infty \lambda \frac{2\alpha\lambda}{(\lambda^2 + \alpha)^2} d\lambda$$

$$= 2\alpha \left[-\frac{\lambda}{2 \cdot (\lambda^2 + \alpha)} + \frac{\arctan \frac{\lambda}{\sqrt{\alpha}}}{2\sqrt{\alpha}} \right]_0^\infty$$

$$= 2\alpha \cdot \frac{\pi}{4\sqrt{\alpha}} = \frac{\sqrt{\alpha}\pi}{2}$$

We can adjust our theoretical distribution because the expectation of the underlying distribution is equal to the expectation of the claim’s distribution.

5.5.2 Cubic case

In this case, we assume:

$$g(\lambda) = \lambda^3$$

so we get:

$$U(\lambda) = \frac{\lambda^3}{\lambda^3 + \alpha}$$

$$u(\lambda) = \frac{3\alpha\lambda^2}{(\lambda + \alpha)^2}$$

For the first order moment, we have:

$$\int \frac{\lambda^3}{\lambda^3 + \alpha} d\lambda = -\frac{\lambda}{3\cdot(\lambda^3 + \alpha)} + \frac{\arctan\left[\frac{-\alpha^{1/3} + 2\lambda}{\sqrt{3}\alpha^{1/3}}\right]}{3\sqrt{3}\alpha^{2/3}}$$
$$+ \frac{\ln\left[\alpha^{1/3} + \lambda\right]}{9\cdot\alpha^{2/3}} - \frac{\ln\left[\alpha^{2/3} - \alpha^{1/3}\cdot\lambda + \lambda^2\right]}{18\alpha^{2/3}}$$

so we get:

$$E[\Lambda] = \int_0^\infty \lambda f(\lambda) d\lambda = \int_0^\infty \lambda \frac{3\alpha\lambda^2}{(\lambda^3 + a)^2} d\lambda$$
$$= 3\alpha\left[\frac{\arctan[\frac{-\alpha^{1/3} + 2\lambda}{\sqrt{3}\alpha^{1/3}}}{3\sqrt{3}\alpha^{2/3}}\right]_0^\infty$$
$$= \alpha\frac{\frac{\pi}{2} + \frac{\pi}{6}}{\sqrt{3}\alpha^{2/3}} = \frac{2\pi}{3\sqrt{3}}\cdot\alpha^{1/3}$$

For the moment of the second order, we have to calculate:

$$\int \frac{\lambda^4}{\lambda^3+\alpha} d\lambda = -\frac{\lambda^2}{3(\lambda^3+\alpha)} + \frac{2\arctan\left[\dfrac{-\alpha^{1/3}+2\lambda}{3\sqrt{3}a^{1/3}}\right]}{3\sqrt{3}\alpha^{1/3}}$$

$$+\frac{2\ln[\alpha^{1/3}+\lambda]}{9\alpha^{1/3}} - \frac{\ln[\alpha^{2/3}-\alpha^{\frac{1}{3}}\cdot\lambda+\lambda^2]}{9\alpha^{1/3}}$$

which gives:

$$E[\Lambda^2] = \int_0^\infty \lambda^2 f(\lambda) d\lambda = \int_0^\infty \lambda^2 \frac{3\alpha\lambda^2}{(\lambda^3+a)^2} d\lambda$$

$$= 3\alpha \left[\frac{2\arctan\left[\dfrac{-\alpha^{1/3}+2\lambda}{3\sqrt{3}a^{1/3}}\right]}{3\sqrt{3}\alpha^{1/3}} \right]_0^\infty$$

$$= 3\alpha \left[\frac{2\dfrac{\pi}{2}}{3\sqrt{3}\alpha^{1/3}} + \frac{2\dfrac{\pi}{6}}{3\sqrt{3}\alpha^{1/3}} \right]$$

$$= \frac{4\sqrt{3}\pi}{9}\alpha^{2/3}$$

Finally, the variance is:

$$VAR[\Lambda] = \frac{4\pi\alpha^{\frac{2}{3}}}{9}(\sqrt{3}-\frac{\pi}{3}).$$

CHAPTER 6:

PRESENTATION OF THE STOCHASTIC VORTICES MODEL AS AN ALTERNATIVE APPROACH TO BONUS MALUS SYSTEMS

6.1 Introduction

In this chapter, we will present the Stochastic Vortices Model based on the assumption that we have an open portfolio, i.e., we consider that a policy can be transferred from one insurance company to another and that the new policies that constantly arrive into a portfolio can be placed not only in the "starting class" but into any of the bonus classes. The Stochastic Vortices Model applies to populations divided into sub-populations which correspond to the transient states of homogeneous Markov chains. Also, by using the limit state probabilities of the Model we can estimate the Long Run Distribution for a BMS and construct optimal BMS scales.

Furthermore, since the Stochastic Vortices Model allows the subscription and the annulment of policies in the portfolio it is an alternative approach to the usual BMS model and the fact that the population is taken as open renders it quite representative of the reality.

6.2 Model Presentation

6.2.1 Stochastic Vortices

As said above, the Stochastic Vortices Model applies to populations divided into sub-populations which correspond to the transient states of homogeneous Markov chains. After a finite time span, every element that has entered the population will end up to an absorbing state and will not belong to the population anymore.

The probabilities of an element that is chosen in the population randomly and belongs to the different sub-populations will be the state probabilities. These state probabilities will converge to limit state probabilities with the use of some quite general conditions. So, the long run distribution will be constituted by the limit state probabilities and it may be used for the study of the performance of the BMS in just the same way as when an alternative model is used.

Notes:

- The state probabilities will be proportional to the mean values of the dimension of the sub-populations.
- In the application to BMS, at first we consider transition probabilities which depend on a parameter λ which will have a structural distribution *U.* Thus firstly, we must obtain a long run distribution that depends on λ and secondly we must decondition it in order to get the unconditional long run distribution which can be used in a performance analysis.

6.2.2 Transition Matrices

As we have already mentioned, apart from *s* transient states, *s* corresponds to the number of the sub-populations, there will be a final absorbing state. Parameter λ will refer to the distribution of the number of claims and in our case we do not have to consider that the transient states constitute a communication class- an assumption that we are led in most of the applications to BMS. The study of the parameter λ in the portfolio will be carried out considering that it was unidimensional but the final deconditioning can also be carried out for a vector of parameters. Finally, we have to mention that transition steps will correspond to years, and that the time $t = 0$ will be the beginning of the year in which the portfolio establishment took place.

For $K_{1,\lambda}$ the *s* x *s* matrix of one step transition between transient states, the full one step transition matrix will be:

$$P_{T,\lambda} = \begin{pmatrix} K_{1,\lambda} & \vec{q}^{\,s}_{1,\lambda} \\ \vec{0}^{s\prime} & 1 \end{pmatrix} \tag{1}$$

the last line corresponds to the absorbing state and the components of $\vec{q}^{\,s}_{1,\lambda}$ are the probabilities for the policyholders in the *s* classes quitting after one year.

<u>Note:</u>

If we add the elements of a row of $K_{1,\lambda}$ with the correspondent component of $\vec{q}^{\,s}_{1,\lambda}$ the sum must be equal to 1.

The above statement points out that in the end of a year, the policyholder either occupies the class foreseen in the transition rules, i.e., he remains in the portfolio or he annuls his policy and does not remain in the company.

<u>Lemma 1</u>

The *n* steps transition matrix will be:

$$P^{(n)}{}_{T,\lambda} = \begin{pmatrix} K_{n,\lambda} & \vec{q}^{\,s}_{n,\lambda} \\ \vec{0}^{s\prime} & 1 \end{pmatrix}$$

with:

$$K_{n,\lambda} = K^{n}_{1,\lambda}$$

$$\vec{q}^{\,s}_{n,\lambda} = \sum_{j=0}^{n-1} K^{j}_{1,\lambda} \cdot \vec{q}^{\,s}_{1,\lambda}$$

Proof:

We have $P^{(n)}{}_{T,\lambda} = P^{n}{}_{T,\lambda}$ because the Markov chain is homogeneous and the proof can be established through mathematical induction once it is observed that:

$$K_{n,\lambda} = K_{1,\lambda} \cdot K_{n-1,\lambda} = K_{1,\lambda} \cdot K_{1,\lambda}^{n-1}$$

$$\vec{q}^{\,s}_{n,\lambda} = K_{1,\lambda} \cdot \vec{q}^{\,s}_{n-1,\lambda} + \vec{q}^{\,s}_{1,\lambda} = K_{1,\lambda} \cdot \sum_{j=0}^{n-2} K_{1,\lambda}^{j} \cdot \vec{q}^{\,s}_{1,\lambda} + \vec{q}^{\,s}_{1,\lambda} = \sum_{j=0}^{n-1} K_{1,\lambda}^{j} \cdot \vec{q}_{1,\lambda}$$

Thus the elements of the matrix $K_{1,\lambda}^{n}$ will be the transition probabilities in n steps between the sub-populations.

If the components of the row vector $\vec{p}^{\,\prime s}_{0,\lambda}$ are the probabilities of a new policyholder being placed in the s classes, the components of:

$$\vec{p}^{\,\prime s}_{n,\lambda} = \vec{p}^{\,\prime s}_{0,\lambda} \cdot K_{1,\lambda}^{n} \tag{2}$$

will be the corresponding probabilities after n years.

6.2.3 Limit state probabilities

With the use of some very general conditions we have (for more see Healy (1986)):

$$K_{1,\lambda} = \sum_{l=1}^{s} \eta_{l,\lambda} \, \vec{\alpha}^{\,s}_{l,\lambda} \, \vec{\beta}^{\,\prime s}_{l,\lambda} \tag{3}$$

, $\eta_{l,\lambda}$ $\left[\vec{\alpha}^{\,s}_{l,\lambda} ; \vec{\beta}^{\,\prime s}_{l,\lambda}\right]; i = 1,\ldots,s$ are the eigenvalues (left and right eigenvectors) of $K_{1,\lambda}$ with:

$$\vec{\beta}'^{s}_{l,\lambda} \cdot \vec{\alpha}^{s}_{h,\lambda} = 0; l \neq h \tag{4}$$

so that

$$K_{n,\lambda} = K^{n}{}_{1,\lambda} = \sum_{l=1}^{s} \eta^{n}{}_{l,\lambda} \vec{\alpha}^{s}_{l,\lambda} \vec{\beta}'^{s}_{l,\lambda} \tag{5}$$

Now (for more see Parzen(1965)), the transition probabilities between transient states tend to zero with the number of steps, thus

$$| n_{l,\lambda} | < 1; l = 1, \ldots\ldots s \tag{6}$$

In the following, will denote as θ_i the mean value of admissions at the *i-th* year, in order to simplify the computation we will assume that the admissions are at the beginning of each year.

<u>Lemma 2</u>

At the end of *n* years, the mean value for the number of policyholders in the different classes will be the components of:

$$\vec{u}'^{s}_{n,\lambda} = \vec{p}'^{s}_{0,\lambda} \cdot \sum_{i=0}^{n} \theta_{n-i} K^{i}_{1,\lambda} \tag{7}$$

<u>Proof:</u>

The sum of the mean values of the policyholders in the different classes that were admitted at years *n* — *i* , *i* = 0,... ,*n.* are mean values we want to find and according to expression (2) these partial mean values will be the components of the row vectors:

$$\theta_{n-i} \cdot \overset{\rightarrow's}{p}_{i,\lambda} = \theta_{n-i} \cdot \overset{\rightarrow's}{p}_{0,\lambda} \cdot K_{1,\lambda}^{i} ; i = 0,.......,n$$

and the lemma is established.

Now we will make an assumption for the growth of the portfolio. Because in most of the cases around the world we have high competitive markets we put:

$$\theta_i = \kappa(1 - e^{-\beta i}), \kappa, \beta \geq 0 \qquad (8)$$

so that the admissions will tend to a limit κ. Finally, even under more aggressive growth limit state probabilities exist because (6) holds.

<u>Proposition</u>

After n years, if (8) holds, the mean values for the number of policyholders in the different classes will be the components of:

$$\overset{\rightarrow's}{u}_{n,\lambda} = \kappa \cdot \overset{\rightarrow's}{p}_{0,\lambda} \cdot \left[\sum_{i=0}^{n} K_{1,\lambda}^{i} - e^{-\beta n} \sum_{i=0}^{n} (e^{\beta} K_{1,\lambda})^{i} \right] \qquad (9)$$

and

$$\overset{\rightarrow's}{u}_{\infty,\lambda} = \lim_{n \to \infty} \overset{\rightarrow's}{u}_{n,\lambda} = \kappa \cdot \overset{\rightarrow's}{p}_{0,\lambda} \cdot \sum_{l=1}^{n} \frac{1}{1 - \eta_{l,\lambda}} \cdot \overset{\rightarrow s}{\alpha}_{l,\lambda} \overset{\rightarrow's}{\beta}_{l,\lambda} \qquad (10)$$

<u>Proof:</u>

The first part of the proposition follows from expressions (7) and (8). Regarding the second part of the proposition, firstly, due to (5) and (6), we have:

$$\sum_{i=0}^{\infty} K_{1,\lambda}^{i} = \sum_{i=0}^{\infty}\sum_{l=1}^{s} \eta_{l,\lambda}^{i} \cdot \overrightarrow{\alpha}^{s}_{l,\lambda} \cdot \overrightarrow{\beta}'^{s}_{l,\lambda} =$$

$$= \sum_{l=1}^{s} \overrightarrow{\alpha}^{s}_{l,\lambda} \cdot \overrightarrow{\beta}'^{s}_{l,\lambda} \sum_{i=0}^{\infty} \eta_{l,\lambda}^{i} = \sum_{l=1}^{s} \overrightarrow{\alpha}^{s}_{l,\lambda} \cdot \overrightarrow{\beta}'^{s}_{l,\lambda} \cdot \frac{1}{1-\eta_{l,\lambda}}$$

and due to

$$\sum_{i=0}^{n} e^{\beta(i-n)} \eta_{l,\lambda}^{i} = e^{-\beta\eta} \frac{1-e^{\beta(\eta+1)}\eta_{l,\lambda}^{n+1}}{1-e^{\beta}\eta_{l,\lambda}} =$$

$$= \frac{e^{-\beta\eta} - e^{\beta}\eta_{l,\lambda}^{n+1}}{1-e^{\beta}\eta_{l,\lambda}} \to 0$$

as $n \to \infty$ for $l = 1,\ldots,s$

we also have:

$$\sum_{i=0}^{n} e^{\beta(i-n)} K_{1,\lambda}^{i} = \sum_{l=1}^{s} \overrightarrow{\alpha}^{s}_{l,\lambda} \cdot \overrightarrow{\beta}'^{s}_{l,\lambda} \sum_{i=0}^{n} e^{\beta(i-n)} \eta_{l,\lambda}^{i} \to 0 \text{ as } n \to \infty \text{ for } l = 1,\ldots,s \tag{11}$$

and the proposition is established.

In order to obtain the limit state probabilities when $\overrightarrow{u}'^{s}_{\infty,\lambda}$ is defined, we have to divide its components by their sum.

Denote as $\overrightarrow{\pi}'^{s}_{T,\lambda}$ the row vector of these limit probabilities which clearly do not depend on *K,* this is, on the upper bound for admissions.

To obtain the unconditional long run distribution we have to decondition, getting:

$$\pi_T(j) = \int_0^{\infty} \pi_{T,\lambda}(j)\, dU(\lambda) \quad , j = 1,\ldots,s \tag{12}$$

, $\pi_{T,\lambda}(j), j = 1,\ldots,s$ are the conditional (unconditional in our case) limit state probabilities.

CHAPTER 7

DESIGN OF AN OPTIMAL BONUS-MALUS SYSTEM WHERE THE CLAIM FREQUENCY DISTRIBUTION IS POISON-INVERSE GAUSSIAN AND THE CLAIM SEVERITY DISTRIBUTION IS PARETO

7.1 Introduction

For the first time in actuarial literature, we will propose a combination of a Poisson-Inverse Gaussian distribution for modeling claim frequency and of a Pareto distribution for modeling claim severity for the construction of an optimal BMS.

At this point we should mention that we choose the Poisson- Inverse Gaussian distribution for modeling claim frequency as an alternative to the Negative Binomial distribution, that is frequently used, based on the fact that mixed Poisson distributions have thicker tails than the Poisson distribution thus they provide a better fit to claim frequency data than the Poisson distribution when the portfolio is heterogeneous. Furthermore, following Frangos and Vrontos (2001) and Mert and Saykan (2005) we choose the heavy-tailed Pareto distribution for modeling claim severity because through the use of long tail distributions apart from many small claim severities, high claim severities can also be observed.

7.2 Design of optimal BMS with a frequency and a severity component based on the a posteriori criteria

Under the basic assumption that the number of claims of each policyholder is independent from the severity of each claim, we propose that the claims number is distributed according to the Poisson- Inverse Gaussian distribution and that the claims severity is distributed according to the Pareto distribution. Using Bayes' theorem, we will find the posterior distributions of the mean claim frequency and the mean claim size, given the information we have about the claim frequency and claim size history for each policyholder. Optimality is obtained by minimizing the insurer's risk.

7.2.1 Frequency component

Under the assumptions that the portfolio is heterogeneous and that all policyholders have constant but unequal underlying risks to have an accident, the number of claims *k*, given the parameter λ,is considered to follow the Poisson(λ) distribution,

$$p_{\lambda}(k \mid \lambda) = \frac{\lambda^{k}}{k!} e^{-\lambda}$$

κ= 0,1,2,3,... and λ > 0 and λ denotes the different underlying risk of each policyholder to have an accident. It is assumed for the structure function that λ ~IG(α, τ) with a probability density function:

$$u(\lambda) = \frac{g}{\sqrt{2\pi h \lambda^{\frac{3}{2}}}} \cdot \exp\left[-\frac{1}{2h\lambda}(\lambda - g)^{2}\right] \text{,with g>0, h>0.}$$

Then it can be proved that the resulting unconditional distribution of the number of claims *k* will be the Poisson- Inverse Gaussian with *E(λ)*= g and *Vα r* (λ) =g(1+h) (for more see Willmot (1986,1987), Venter (1991a), Besson and Patrat (1992), Trembley (1992) and Lemaire (1992)).The probabilities P(*k*) can be calculated recursively:

$$P(0) = e^{\frac{g}{h}\left[1-(1+2h)^{\frac{1}{2}}\right]}$$

$$P(1) = gP(0)(1+2h)^{-\frac{1}{2}}$$

$$(1+2h)k(k-1)P(k) = h(k-1)(2k-3)P(k-1) + g^{2}P(k-2), k = 2,3,...$$

The moments estimators of $\hat{g}, \hat{h}$ are: $\hat{g} = x$ and $\hat{h} = \frac{s^{2}}{x} - 1$ (providing $s^{2} > x$).

For K=$\sum_{i=1}^{t} k_i$ the total number of claims that a policyholder had in t years, with k_i the number of claims that the policyholder had in the year i, i = 1,...,t using the Bayes' theorem we will obtain the posterior structure function of λ for a policyholder with claim history k_1,....k_t,(for more see Besson and Partrat (1992) and Tremblay (1992)).

$$u(\lambda \mid \kappa_1,\ldots,\kappa_t) = \frac{P(\kappa_1,\ldots,\kappa_t \mid \lambda)u(\lambda)}{\int_0^\infty P(\kappa_1,\ldots,\kappa_t \mid \lambda)u(\lambda)d\lambda} =$$

$$= \frac{\dfrac{\lambda^k e^{-t\lambda}}{\prod (k_j!)} \dfrac{g}{\sqrt{2\pi h \lambda^{-\frac{3}{2}}}} \exp\left[-\dfrac{1}{2h\lambda}(\lambda - g)^2\right]}{\int_0^\infty \dfrac{\lambda^k e^{-t\lambda}}{\prod (k_j!)} \dfrac{g}{\sqrt{2\pi h \lambda^{-\frac{3}{2}}}} \exp\left[-\dfrac{1}{2h\lambda}(\lambda - g)^2\right] d\lambda}$$

After simplification of g, $\Pi(k_j!)$ and $\sqrt{2\pi h}$ and deletion of the integral as a normalizing constant:

$$u(\lambda \mid \kappa_1,\ldots,\kappa_t) \propto \frac{\lambda^k e^{-t\lambda}}{\lambda^{-3/2}} \exp\left[-\frac{1}{2h\lambda}(\lambda - g)^2\right]$$

$$\propto \lambda^{k-\frac{3}{2}} e^{-\lambda(t+\frac{1}{2h})} e^{-\frac{1}{\lambda}(\frac{g^2}{2h})} = \lambda^{\alpha} e^{-\frac{\lambda}{b}} e^{-\frac{c}{\lambda}}$$

setting $a = k - \frac{3}{2}$, $\frac{1}{b} = t + \frac{1}{h}$, $c = \frac{g^2}{2h}$.

With $\mathrm{a} = u - 1$, $b = 2\beta$ and $c = \mu^2 / 2\beta$, this is the density function of a generalized inverse Gaussian distribution, defined in most cases as:

$$f(\lambda)=\frac{\lambda^{v-1}e^{-\frac{\lambda}{2\beta}}e^{-\frac{\mu^2}{2\beta\lambda}}}{2\mu^{v}\mathrm{K}_v(\frac{\mu}{\beta})}$$

K_v, is the modified Bessel function of the third kind with index v, defined as:

$$\mathrm{K}_v(u)=\frac{1}{2}\int_0^{\infty}x^{v-1}e^{-\frac{u}{2}(x+\frac{1}{x})}dx$$

For all u>0, $\mathrm{K}_v(u)$ satisfies the two properties:

$$\mathrm{K}_{-v}(u)=\mathrm{K}_v(u)$$

$$\mathrm{K}_{v+1}(u)=\frac{2v}{u}\mathrm{K}_v(u)+\mathrm{K}_{v-1}(u)$$

The generalized-inverse Gaussian distribution is a conjugate family for the Poisson, it reduces to inverse Gaussian when u=-1/2. The average of the generalized-inverse Gaussian distribution is:

$$\mu\frac{\mathrm{K}_{v+1}(\frac{\mu}{\beta})}{\mathrm{K}_v(\frac{\mu}{\beta})}=\mu\frac{\mathrm{K}_{k+\frac{1}{2}}(\frac{\mu}{\beta})}{\mathrm{K}_{k-\frac{1}{2}}(\frac{\mu}{\beta})}$$

Denote:

$$Q_k(u)=\frac{\mathrm{K}_{k+\frac{1}{2}}(u)}{\mathrm{K}_{k-\frac{1}{2}}(u)}$$

Using the quadratic error loss function the optimal choice of λ_{t+1} for a policyholder with claim history $k_1,....,k_t$ will be the mean of the posterior structure function. According to the above calculations we find that:

$$\lambda_{t+1}(k_1,...,k_t) = \mu Q_k(\frac{\mu}{\beta}) \qquad (1)$$

7.2.2 Severity component

We will denote as x the size of the claim of each insured and as y the mean claim size for each insured and we will assume that the conditional distribution of the size of each claim given the mean claim size $x|y$, for each policyholder is the exponential distribution with parameter y with a probability density function given by:

$$\mathrm{f(x|y)} = \frac{1}{\mathrm{y}} \cdot e^{-\frac{x}{y}}$$

for $x > 0$ and $y > 0$. Also $E(x|\ y)=y$ and $var(x\ y) = y^2$. The mean claim size y is not the same for all the policyholders so, the prior distribution of y is IG(s, m) and it's probability density function is given by:

$$g(y) = \frac{\frac{1}{m} \cdot e^{-\frac{m}{y}}}{\left(\frac{y}{m}\right)^{s+1} \cdot \Gamma(s)}$$

The expected value of the mean claim size y is:

$$E(y) = \frac{m}{s-1}$$

The unconditional distribution of the claim size x is:

$$P(X=x)=\int_0^{\infty} f(x|y)\cdot g(y)dy =$$
$$= s\cdot m^{s}\cdot\left(x+m\right)^{-s-1}$$

the probability density of the Pareto distribution with parameters s and m. Thus Pareto distribution can be generated in the following way: if $x|y \sim Exponential(y)$ and $y \sim$ Inverse $Gamma(s,m)$ then $x \sim Pareto(s,m)$. So the relatively tame exponential distribution gets transformed in the heavy-tailed Pareto distribution, which is good for modeling the claim severity instead of the exponential distribution which is often inappropriate. Taking y distributed according to the Inverse Gamma, heterogeneity that characterizes the severity of the claims of different policyholders is incorporated in the model.

Then we will calculate the posterior distribution $g(y|x_1, \dots, x_k)$ in order to design an optimal BMS that will take into account the claim severity. Denote for a policyholder that was in the portfolio for t years:

1. $K=\sum_{i=1}^{t} k_i$ the total number of claims that he had

in t years with k_i the number of claims that the policyholder had in the year i, $i = 1,\dots,t$.

2. The vector $x_1, x_2, \dots, x_k$ his claim size history and $\sum_{k=1}^{K} X_k$ his total claim amount.

Applying Bayes' theorem:

$$g(y \mid x_1,\ldots,x_k) = \frac{\dfrac{1}{\left(m+\sum_{k=1}^{K} x_k\right)} \cdot e^{-\frac{m+\sum_{k=1}^{K} x_k}{y}}}{\left(\dfrac{y}{m+\sum_{k=1}^{K} x_k}\right)^{K+s+1} \cdot \Gamma(K+s)}$$

the posterior distribution of the mean claim size y given the claims size history of the policyholder x_1, ..., x_k and it is the Inverse Gamma($s + K$, m+$\sum_{k=1}^{K} x_k$). This update of the parameters of the Inverse Gamma from s and m to s+K *and* m+$\sum_{k=1}^{K} x_k$ is necessary due to the occurrence of K claims in t years with aggregate claim amount equal to $\sum_{k=1}^{K} x_k$. Also Inverse Gamma distribution is said to have the important property of being conjugate with the exponential likelihood. Finally, the mean of the posterior distribution of the mean claim size is:

$$E(x \mid y) = \frac{m+\sum_{k=1}^{K} x_k}{s+K-1} \tag{2}$$

and the predictive distribution of the size of the claim of each insured x will be also a member of the Pareto family.

7.2.3 Calculation of the Premium according the Net Premium Principle

Consider a group of policyholders who stayed in the portfolio for t years and had K

claims with $\sum_{k=1}^{K} x_k$ total claim amount. We have obtained the expected number of claims $\lambda_{t+1}(k_1, ..., k_t)$ by (1) and the expected claim severity $y_{t+1}(x_1, ..., x_K)$ by (2).Thus the net premium that must be paid from that group, according the net premium principle is given by :

$$Premium = P_{t+1}(k_1,.....k_t) = \mu Q_k(\frac{\mu}{\beta}) \cdot \frac{m + \sum_{k=1}^{K} x_k}{s + K - 1}.$$

Furthermore, the two properties of $\mathrm{K}_v(u)$ imply that:

$$Q_0(u) = 1$$

$$Q_k(u) = \frac{2k-1}{u} + \frac{1}{Q_{k-1}(u)}$$

Thus, we can avoid numerical integration and enable a recursive calculation of the optimal BMS. Also in order to find the net premium we have to know:

- g and h , the parameters of the Poisson- Inverse Gaussian distribution (see Lemaire (1995) for their estimation)
- s and m the parameters of the Pareto distribution (see Hogg and Klugman (1984) for their estimation)
- the number of years t that the policyholder was in the portfolio,
- his number of claims K and his total claim amount $\sum_{k=1}^{K} x_k$

CHAPTER 8

CONCLUSIONS

The objectives of the current thesis were to study optimal bonus-malus systems in automobile insurance using different underlying approaches. As we have mentioned in several occasions, the majority of optimal bonus-malus systems assign to each policyholder a premium based on the number of his accidents disregarding their size. In this way, a policyholder who underwent an accident with a small size of loss will be unfairly penalized in the same way with a policyholder who had an accident with a big size of loss.

In chapter 2, following the work of Frangos and Vrontos (2001) we presented an optimal BMS based on the a posteriori frequency and the a posteriori severity component. More specifically, Frangos and Vrontos have expanded the classical, in the BMS literature, model of Lemaire (1995) who used the Negative Binomial distribution. The construction took place under the assumption that the number of claims was distributed according to the Negative Binomial distribution and that the losses of the claims were distributed according to the Pareto distribution. By the application of Bayes' theorem, the posterior distributions of the mean claim frequency and the mean claim size were calculated given the known claim frequency history and the claim size history for each policyholder, for all the time that he/she stayed in the portfolio (for more information on this subject, refer to Vrontos (1998)). Optimality was obtained by minimizing the insurer's risk and the optimal BMS had all the attractive properties of the one designed by Lemaire.

Furthermore, important a priori information for each policyholder was incorporated in this design of optimal BMS and a generalized BMS based both on the a priori and a posteriori classification criteria was proposed. In the generalized BMS, the premium was a function of the years that the policyholder was in the portfolio, of his claim number, of his aggregate claim amount, of the statistically significant a priori rating variables for the number of accidents and for the size of loss that each of these claims incurred. In this case Frangos and Vrontos have extended the model developed by Dionne and Vanasse (1989, 1992), which was based only on the frequency component, by taking into consideration the individual's characteristics, the number of his accidents and by introducing a severity component and the generalized optimal BMS had all the attractive properties of the one obtained by Dionne and Vanasse (1989, 1992).

Finally, Frangos and Vrontos used for their application a data set of a Greek insurance company (for more refer to Frangos and Vrontos (2001)) and presented the following conclusions:

- The basic advantage of the optimal BMS based on the a posteriori frequency and severity component in comparison with the one that takes under consideration only the frequency component is the differentiation of the premiums according to the severity of the claim.
- The generalized optimal BMS with a frequency and a severity component based both on the a priori and the a posteriori classification criteria, in comparison with the one obtained when only the a posteriori frequency and severity component are used, is more fair since it considers all the important a priori and a posteriori information for each policyholder both for the frequency and the severity component and thus it permits the differentiation of the premiums for various number of claims and for various claim amounts based on the estimations of the expected claim frequency and the expected claim severity of each policyholder.

In chapter 3, we presented an optimal BMS that takes into account the claim frequency and one that takes into account both the claim frequency and the claim severity based on the work of Mert and Saykan (2005). A mixture of a Poisson and Exponential distributions, the Geometric distribution was used for modeling claim numbers and the Pareto distribution was used again, following Frangos and Vrontos (2001), for modeling claim amounts. The risk premium was calculated using the net premium principle, and the results were obtained by using the claim number only and by using both the claim number and claim amount.

In Mert's and Saykan's application the two BMS were compared and, as in the work of Frangos and Vrontos (2001), it was concluded that the optimal BMS based on the claim frequency and claim severity component is fairer for the policyholders than the optimal BMS based only on the claim frequency component because it differentiates the premiums (for more refer to Mert and Saykan (2005)). Nevertheless, Mert's and Saykan's (2005) design of optimal BMS was based only on the a posteriori information they had about every policyholder and the a priori information was neglected. Thus they didn't propose a generalized optimal BMS like Frangos and Vrontos(2001). The expansion of their design for the construction of a generalized optimal BMS based both on the a priori and a posteriori classification criterion using the same claim frequency and claim severity distributions is an interesting topic for further research.

In chapter 4 based on the work of Walhin and Paris (1997) we constructed an optimal BMS using as the claim frequency distribution a three parameters distribution encompassing the Negative Binomial and the Poisson-Inverse Gaussian the Hofmann's distribution. The net premium principle or the principle of zero utility was used for the construction. The formulae were easily derived using the Bayes theorem and the form of the mixed Poisson distribution. Furthermore, a non-parametric method, that permits a simple formulation of the stationary and transition probabilities in a portfolio, was proposed for the construction of an optimal BMS.

In Walhin's and Paris's application, we noticed that the parametric mixed Poisson distribution used is more general than the traditional Negative Binomial or the Poisson Inverse Gaussian distributions but it has the disadvantage that three parameters have to be estimated. However, one parameter for the maximum likelihood estimation is given by the experimental mean. Also, the fit is slightly better than the Negative Binomial or the Poisson Inverse Gaussian fits. About the non-parametric case, we noticed that it is better than the three parameters parametric distribution in terms of goodness of fit. The parameters can be estimated easily because it has a simple form.

Nevertheless, we have to mention that the parametric approach should be preferred for the construction of a BMS due to its 'continuity'. Actually, a nice form is given to the bonus-malus table by the continuous form of the distribution of Λ while the discontinuity of the non-parametric distribution gives the bonus-malus table a curious comportment with local almost constant premiums. Furthermore, we can use the tables constructed with the parametric distribution as a starting point for methods like the one proposed by Coene and Doray (1996) to 'fit' bonus-malus tables with classes.

The non-parametric fit should be preferred for the evaluation of the mean of difficult functions of λ over the portfolio as it can replace the complicated numerical integration by a short summation. An example of the advantage of the non-parametric fit in a portfolio of a BMS with finite number of classes is the calculation of the stationary and transient distributions. Finally, by the use of the non-parametric fit we can easily evaluate other quantities like the mean asymptotic efficiency of Loimaranta (1972), (for more details refer to Walhin and Paris (1997)).

In chapter 5, our analysis was based on the fact that for the construction of optimal BMS the distribution of the number of car accidents is frequently chosen within the "mixed-Poisson" family. Following the work of Hulin and Justens (Banque Pictet-Luxembourg HEFF/Cooremans Bruxelles) we showed the general properties of "mixed Poisson" family distributions and we gave a unifying approach of several particular cases including the geometrical, the P-Erlang, the Negative Binomial and the P-inverse Gaussian distributions. Nevertheless, the main problem in the data fitting remained the thickness of the distribution tail and the systematic use of a negative exponential never led to a sufficiently large opportunity for a great number of claims to occur. Thus, Hulin and Justens introduced a new family of "mixed-Poisson", built upon “fatty-tailed” underlying distributions, the “P-rational’’ distributions in order to avoid the problem of adjustment that is the thickness of the tails of the underlying distributions.

In their applications Hulin and Justens, adjusted the common distributions to several data bases: the Belgian, the French and the Italian. and showed the weakness of the classical models, needing the regrouping of the worst cases and working with 2 parameters at least (for a distribution taking 5, 6 or 7 values, which means 4 to 6 degrees of freedom without adjustment). Also, graphical comparisons between the “P-rational distributions” and the classical distributions showed that the tails of the rational underlying distributions had the desirable properties and. The “P-rational distributions fitted the data, even if some of them appeared to have an infinite variance, which led to a rather astonishing problem: distributions with infinite theoretical variance may sometimes better fit observed data. Thus, the existence of the variance is not necessarily needed in practical problems (for more details refer to Hulin and Justens (Banque Pictet-Luxembourg HEFF/Cooremans Bruxelles)).

In chapter 6, we presented an alternative approach to BMS the Stochastic Vortices Model based on the work of Guerreiro and Mexia (2004). The Stochastic Vortices Model was developed under the assumption that we had an open portfolio, i.e., it was considered that a policy could be transferred from one insurance company to another and that the new policies that constantly arrived into the portfolio could be placed not only in the “starting class” but into any of the bonus classes. The Stochastic Vortices Model is applicable to populations divided into sub-populations which correspond to the transient states of homogeneous Markov chains. Also, the Long Run Distribution for a BMS can be estimated by the use of the limit state probabilities of the Model and optimal bonus-malus scales can be easily calculated.

For their application, Guerreiro and Mexia used a data set of a recently established Portuguese insurance company. They applied the Stochastic Vortices Model to this data set obtaining the correspondent long run distribution. Afterwards, some optimal bonus scales, such as Norberg’s [1979], Borgan, Hoem & Norberg’s [1981], Gilde & Sundt’s [1989] and Andrade e Silva’s [1991] were estimated. Furthermore, they compared their results with the classic model (Closed Model) for BMS and with the Open Model developed by Centeno & Andrade e Silva [2001]. The Stochastic Vortices Model and the Open Model had quite similar optimal bonus scales, avoiding the very low or very high premiums that were obtained when using the closed model. Thus, the two first models were preferred to the closed model which over evaluated the probabilities for the higher classes since it did not took into consideration the policyholders tendency to leave when they attain higher maluses. The Stochastic Vortices Model and the Open Model were the more realistic ones (for more details refer to Guerreiro and Mexia).

At this point it should be mentioned that in Portugal, the transfer of information between insurers is not efficient. Thus, policyholders who had a claim during an annuity, leave their insurance company in order to buy another policy to a competitor declaring that this is their first insurance policy. In this way, they are not penalized and they can be treated by their new insurer as individuals who never had a claim. Also, it should not be assumed that all policyholders start in the same class. That is because due to commercial goals, discounts are given to new policyholders thus they are not all placed in the pre-defined “starting class”. Also, in other cases, when the insurance company requests the Tariff Certificate, the new policyholders start in an aggravated class. Thus, the use of Stochastic Vortices Model in the application of Guerreiro and Mexia allowed the subscription and the annulment of policies in the portfolio as it was an alternative approach to the usual BMS model and the fact that the population is taken as open renders it quite representative of the reality.

Furthermore, in situations that correspond to open populations in which there are entrances and departures the irreducible Markov chains of discrete parameter fail and the Stochastic Vortices offer useful models. When the population is divided into sub-populations and the transition probabilities are invariant between sub-populations, these sub-populations can be considered as the transient states of a homogenous Markov chain. Under quite general assumptions in the entrances, the limit state probabilities correspond to the long run distribution can be obtained. Also, because the population is taken as open, this kind of models is much more realistic than classic models based in a closed population occupying the states of an irreducible Markov chain.

In chapter 7 for the first time in actuarial literature we proposed a combination of a Poisson- Inverse Gaussian distribution for modeling claim frequency and of a Pareto distribution for modeling claim severity for the construction of an optimal BMS. We chose the Poisson- Inverse Gaussian distribution for modeling claim frequency as an alternative to the Negative Binomial distribution, that is frequently used, based on the fact that mixed Poisson distributions have thicker tails than the Poisson distribution. Thus they provide a better fit to claim frequency data than the Poisson distribution when the portfolio is heterogeneous. Furthermore, following Frangos and Vrontos (2001) and Mert and Saykan (2005) we choose the heavy-tailed Pareto distribution for modeling claim severity because through the use of long tail distributions, apart from many small claim severities, high claim severities can also be observed.

Finally, an interesting topic for further research could be the expansion of the above design for the construction of a generalized optimal BMS based both on the a priori and the a posteriori classification criteria using the same claim frequency and claim severity distributions or the design of optimal BMS with different claim frequency and claim severity distributions.

REFERENCES

Albrecht, P. (1980). "On the Correct Use of the Chi-Square Goodness-of -Fit Test." *Scandinavian Actuarial Journal. 149-170. Comment in Scandinavian Actuarial Journal 1982\ 168-170.*

Albrecht, P. (1982a). "On Some Statistical Methods Connected With the Mixed Poisson Distribution." *Scandinavian Actuarial Journal. 1-14.*

Albrecht, P. (1982b). "Testing the Goodness of Fit of a Mixed Poisson Process." *Insurance: Mathematics and Economics. 1, 27-33.*

Baxter, L.A., Coutts S.M. and Ross G.A.F. (1979). "Applications of Linear Models in Motor Insurance. " *21st International Congress of Actuaries.*

Beirlant, J., V. Derveaux, A.M. de Meyer, M. Goovaerts, E. Labie, and B. Maenhout. (1991). "Statistical Risk Evaluation Applied to (Belgian) Car Insurance." *Insurance: Mathematics and Economics. 10, 289-302.*

Besson, J.-L., and C. Partrat. (1992). "Trend et systemes de bonus-maius." *ASTIN Bulletin. 22. 11-31.*

Bichsel, F. (1960). "Une methode pour calculer une ristoume adequate pour annees sans sinistres." *ASTIN Bulletin. 1, 106-112.*

Bichsel, F. (1964). "Erfahrung-Tarifierung in der Motorfahrzeug-haftplichtversicherung." *Mitteiluneen der Vereinigung Schweizerischer Versicherungsmathematiker. 119-129.*

Bonsdorff, H. (1992). "On the Convergence Rate of Bonus-Malus Systems." *ASTIN Bulletin. 22, 217-223.*

Borgan, 0. J. Hoem, and R. Norberg. (1981). "A Non Asymptotic Criterion for the Evaluation of Automobile Bonus Systems." *Scandinavian Actuarial Journal. 165-178.*

Boulanger, F. (1994). "Systeme de Bonus-MaJus multi-garanties." *ASTIN*

Colloquium, Cannes.

Brockman, M.S., and T.S. Wright. (1992). "Statistical Motor Rating: Making Effective Use of Your Data." *Journal of the Institute of Actuaries. 119 III.*

Buhlmann, H. (1964). "Optimale Pramienstufensysteme." Mitteiluneen der Vereiniaung Schweizerischer Versicherungsmathematiker. 193-213.

Buhlmann, H. (1970). "Mathematical Methods in Risk Theory. " *Berlin: Springer.*

Centeno, L. e Andrade e Silva, J. (2001). "Bonus Systems in Open Portfolio. " *Insurance Mathematics e Economics, pp. 341-350.*

Coene, G. and Doray, L.G. (1996). "A Financially Balanced Bonus-Malus System. " *ASTIN Bulletin 26, 107-115.*

Consul, P.C. (1990). "A Model for Distributions of Injuries in Auto Accidents." *Mitteilungen der Vereinieung Schweizerischer Versicherungsmathematiker. 163-168.*

Corlier, F., J. Lemaire, and D. Muhokoio. (1979c). "Simulation of an Automobile Portfolio." *Geneva Papers on Risk and Insurance. 40-46.*

Cousy, H. and H. Claassens. (1994). "Ex Post Control of Insurance Policies in Belgium." *Geneva Papers on Risk and Insurance. 70, 46-59.*

De Leve, G. and P.J. Weeda (1968). "Driving with Markov-Programming." *ASTIN Bulletin, 5, 62-86.*

De Pril, N. (1978). "The Efficiency of a Bonus-Malus System. " *ASTIN Bulletin, Vol. 10 Part 1, pp. 59-72.*

De Pril, N. (1979). "Optimal Claim Decisions for a Bonus-Malus System: A Continuous Approach." *ASTIN Bulletin, 10, 215-222.*

De Pril, N., and M. Goovaerts. (1983). "Bounds for the Optima! Critical Claim Size of a Bonus System." *Insurance: Mathematics and Economics. 2, 27-32.*

Delaporte, P. (1960). "Un probleme de tarification de I'assurance accidents d'automobiles examine par la statistique mathematique." *Proceedings of the Sixteenth*

International Congress of Actuaries, Bruxelles, 2, 121-135.

Delaporte, P. (1965). "Tarification du risque individuel d'accidents automobiles par la prime modeiee sur le risque." *ASTIN Bulletin, 3, 251-271.*

Delaporte, P. (1972a). "Les mathematiques de I'assurance automobile." *ASTIN Bulletin. 6, 185-190.*

Delaporte, P. (1972b). "Construction d'un tarif d'assurance automobile base sur le principe de Ja prime modeiee sur le risque." *Mitteitungen der Vereinigung Schweizeriseher Versicherungsmathematiker. 101-113.*

Dellaert, N.P., JBG Frenk, A. Kouwenhoven, and B.S. Van der Laan. (1990). "Optimal Claim Behaviour for Third-Party Liability Insurances or to Claim or Not to Claim: That Is the Question." *Insurance: Mathematics and Economics, 9, 59-76.*

Dellaert, N.P., JBG Frenk, and E. Voshol. (1991). "Optimal Claim Behaviour for Third-Party Liability Insurances with Perfect Information." *Insurance: Mathematics and Economics, 10, 145-151.*

Depauw, J., M. De Rouck, F. Delbaen, J. Haezendonck, and P. Van Goethem (1980). "Etude mathematique de I'assurance automobile beige." *Bulletin de PAssociation Rovale des Actuaires Beiges, 21-68.*

Dionne, G., and C. Vanasse (1989). "A Generalization of Automobile Insurance Rating Models: The Negative Binomial Distribution with a Regression Component" *ASTIN Bulletin. 19, 199-212.*

Dionne, G., and C. Vanasse (1992). "Automobile insurance ratemaking in the presence of asymmetrical information. " *Journal of Applied Econometrics 1, 149-165.*

Dufresne, F. (1988). "Distributions stationnaires d'un systeme bonus-malus et probability de ruine." *ASTIN Bulletin. 18, 31-46.*

Dufresne, F. (1995). "The Efficiency of the Swiss Bonus-Malus System." *Bulletin of the Swiss Actuaries, 1995(1) 29-41.*

Dureuil, G., and C. Geoffrey. (1994). "Modelisation bivariee de frequences de sinistres dependantes." *ASTIN Colloquium, Cannes.*

Feller, W. (1971). "An Introduction to Probability Theory and its Applications. " *New York: Wiley.*

Ferreira, J. (1974). "The Long-term Effect of Merit-Rating Plans on Individual Motorists." *Operations Research. 22, 954-978.*

Ferreira, J. (1977). "Identifying Equitable Insurance Premiums for Risk Classes: An Alternative to the Classical Approach. " *Twenty-third International Meeting of the Institute of Management Sciences, Athens.*

Frangos, N. and Vrontos, S. (2001). " Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance. " *ASTIN Bulletin, Vol. 31, No.1, pp. 1-22.*

Geller, H. (1979). " An introduction Mathematical Risk Theory. " *University of Pennsylvania.*

Gilde, V., and B. Sundt. (1989a). "On Bonus Systems with Credibility Scales." *Scandinavian Actuarial Journal. 13-22.*

Goovaerts, M., De Vijlder, F. and Haezendonck, J. (1984). "Insurance Premiums: Theory and Applications. " *Amsterdam: North Holland.*

Gourieroux, C., Montfort, A. and Trongoton. A. (1984 a.) "Pseudo maximum likelihood methods: theory." *Econometrica, 52, 681-700.*

Gourieroux, C., Montfort, A. and Trongoton. A. (1984 b.) " Pseudo maximum likelihood methods: application to Poisson models. " *Econometrica, 52, 701-720.*

Guerreiro, G. R. and Mexia, J. T. (2004). "An alternative approach to bonus-malus. "*Discuss. Math. Probab. Stat.24, no.2.*

Gregorius, F. (1982). "Development of the Study." *In New Motor Rating Structure in* *the Netherlands. ASTIN Groep Nederland, 15-37.*

Gurtler, M. (1963). "Bonus ou malus." *ASTIN Bulletin. 3, 43-61.*

Hausmann J.A., Hall B.H. and Griuches Z. (1984). "Econometric models for count data with an application to the patents-R&D relationship. " *Econometrica, 46, 1251-1271.*

Healy, M. (1986). "Matrices for Statistics. " *Oxford Science Publications.*

Herzog, T. (1996). "Introduction to Credibility Theory. " *Actex Publications, Winstead.*

Hogg, R.V. and Klugman, S.A. (1984). " Loss Distributions. "*John Wiley & Sons, New York.*

Hoi-Man, M. (1955). " Uber zusammengesetzte Poiisson-prozesse und thre Anwendungen in der Unfall versicherung. " *Bulletin of the Swiss Actuaries, 55 499-575.*

Holtan, J. (1994). "Bonus Made Easy". *ASTIN Bulletin. 24, 61-74.*

Hulin, L (1999). "Modelisation de la survenance d'accidents." G.E.M.M.E, 9924, Universite de Liege.

Hulin, L. and Justens, D. "General properties of mixed Poisson distributions application to data fitting in the case of ‘automobile collision claims. " *Banque Pictet-Luxembourg HEFF/Cooremans Bruxelles.*

Hurlimann, W. (1990). "On Maximum Likelihood Estimation for Count Data

Models." *Insurance Mathematics and Economics, 9 39-49.*

Islam, M.N. and P.C. Consul (1992). "A Probabilistic Model for Automobile Claims." *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker. 85-93.*

Justens, D. (1996). "Construction interpreted de distributions de probability a valeurs dans les naturels. " *G.E.M.M.E, 9603, Universite de Liege.*

Kalb, G.R.J. - Kofman, P. - Vorst, T.C.F. (1996). "Mixtures of tails in clustered automobile collision claims. " *Insurance, Mathematics and Economics, 18, pp 89-107.*

Kestmont, R. M. and Paris, J. (1985). "Sur la Ajustement du Nombre de Sinistres." *Bulletin of the Swiss Actuaries, 85 157-164.*

Lawless, J.F. (1987). "Negative Binomial Distribution and Mixed Poisson Regression. " *Canadian Journal of Statistics, 15, 3, 209-225.*

Lemaire, J. (1975). "Si les assures connaissaient la programmation dynamique." *Bulletin de la Association Rovale des Actuaires Beiges. 54-63.*

Lemaire, J. (1976). "Driver Versus Company: Optimal Behaviour of the Policyholder." *Scandinavian Actuarial Journal. 209-219*.

Lemaire, J. (1977). "Critique du tariff automobile responsablite civil belge." *Association Rovale des Actuaires Beiges, pp. 93-109.*

Lemaire, J. (1977a). "Lasoif du bonus." *ASTIN Bulletin. 9, 181-190.*

Lemaire, J. (1977b). "Selection Procedures of Regression Analysis Applied to Automobile Insurance." *Mitteilungen der Vereinigunq Schweizerischer Versiefaerengsroatheroatiker. Part I. 1977, 143-160.*

Lemaire, J. (1977c). "Critique du tarif automobile beige". *Bulletin de i'Association*

Rovale des Actuaires Beiges. 93-109.

Lemaire, J. (1979a). "Selection Procedures of Regression Analysis Applied to Automobile Insurance." *Mitteilungen der Vereinieung Schweizerischer Versieherungsmathematiker. Part II, 1979, 65-72.*

Lemaire, J. (1979b). "How to Define a Bonus-Malus System With an Exponential Utility Function." *ASTIN Bulletin. 10, 274-282.*

Lemaire, J. (1984). "The Effect of Expense Loadings on the Fairness of a Tariff." *ASTIN Bulletin. 14, 165-171.*

Lemaire, J. (1984). "Automobile insurance: Actuarial models." *Kluwer-Nijhoff, Netherlands.*

Lemaire, J. (1988a). "Construction of the New Belgian Motor Third Party Tariff Structure." *ASTIN Bulletin. 18. 99-11?*

Lemaire, J. (1988b). "A Comparative Analysis of Most European and Japanese Bonus-Malus Systems." *Journal of Risk and Insurance. 55, 660-681.*

Lemaire, J. (1992). "Negative Binomial or Poisson-Inverse Gaussian?" *Proceedings of the Twenty-fourth International Congress of Actuaries. Montreal.*

Lemaire, J. (1993). "Selecting a Fitting Distribution for Taiwanese Automobile Losses." *Unpublished manuscript.*

Lemaire, J., and H. Zi. (1994a). "High Deductibles Instead of Bonus-Malus: Can it Work?" *ASTIN Bulletin. 24, 75-86.*

Lemaire, J., and H. Zi. (1994b). "A Comparative Analysis of 30 Bonus-Malus Systems". *ASTIN Bulletin. 24, 287-309.*

Lemaire, J. (1995). "Bonus-Malus Systems in Automobile Insurance." *Kluwer Academic Publishers.*

Lindsay, B. (1995). "Mixture Models Theory, Geometry and Applications." *Pennsylvania State University.*

Loimaranta, K. (1972). "Some Asymptotic Properties of Bonus Systems." *ASTIN Bulletin. 6, 233-245.*

Maceda. E. (1948). "On the Compound and Generalized Poisson Distributions."*Annals of Mathematical Statistics, 19, 414-416.*

Mack T. (1991). "A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserves." *ASTIN Bulletin., vol. 21, no.l, pp. 93-109.*

Martin, D.B. (1960). "Automobile Insurance: Canadian Accident-Free Classification System." *ASTIN Bulletin. 1, 123-133.*

Martin-Lof, A. (1973). "A Method for Finding the Optimal Deductible Rule for a Policyhoider of an Insurance Company With a Bonus System." *Scandinavian Actuarial Journal. 23-29.*

McCulagh, P. and Nelder, J.A. (1983, 1989). " Generalized Linear Models. " *Chapman & Hall.*

Mert, M. and Saykan, Y. (2005). "On a bonus-malus system where the claim frequency distribution is Geometric and the claim severity distribution is Pareto. " *Hacettepe Journal of Mathematics and Statistics Vol. 34, pp.75-81.*

Molnar, D., and T. Rockwell. (1966). "Analysis of Policy Movement in z Merit-Rating Program: An Application of Markov Processes." *Journal of Risk and Insurance. 265-276.*

Murray, M.L. (1971). "A Deductible Selection Mode!: Development and Applications." *Journal of Risk and Insurance. 38, 423-436.*

Neuhaus, W. (1988). "A Bonus-Malus System in Automobile Insurance." *Insurance: Mathematics and Economics. 7, 103- i J2.*

Nonnan, J.M., and D.C.S. Shearns. (1980). "Optimal Claiming on Vehicle Insurance Revisited." *Journal of the Operational Research Society. 31, 181-186.*

Norberg, R. (1976). "A Credibility Theory for Automobile Bonus Systems." *Scandinavian Actuarial Journal. 92-107.*

Ortiz, E. (1990). "A Stochastic Model of the Distribution of Wars in Time." *American Statistical Association Annual Meeting. Los Angeles.*

Ottaviani, G. "Sur les criteres de choix des assures dans l'assurance R.C.A." *Conference prononcee au X IV congres ASTIN.*

Panjer, H.H. (1987). "Models of Claim Frequency." *In Advances in the Statistical Sciences: Actuarial Science. VI, 115-125. LB. MacNeill and G J. Umphrey (eds.).*

Panjer, H.H., and G.E. Willmot. (1992). "Insurance Risk Models". *Schaumburg, III.: Society of Actuaries.*

Panjer, H.H., and G.E. Wtllmot. (1988a). "Motivating Claims Frequency Models." *Proceedings of the Twenty-third International Congress of Actuaries, Helsinki, 3, 269-284.*

Partrat, C. (1993). "Compound Poisson Models for Two Types of Claims." *ASTIN Colloquium, Cambridge.*

Parzen, E. (1965). "Stochastic Processes. " *Holden Day, S. Francisco.*

Pashigian, B.P., L.L. Schkade, and G.H. Menefee. (1966). "The Selection of an Optimal Deductible for a Given Insurance Policy." *Journal of Business, 39, 35-44.*

Pesonen, E. (1962). "A Numerical Method of Finding a Suitable Bonus Scale." *ASTIN Bulletin. 2, 102-108.*

Picard, P. (1976). "Generalisation de P etude sur la survenance des sinistres en assurance automobile." *Bulletin Trimestriel de PInstitut des Actuaires Francais. 204-*

267.

Picech, L. (1994). "The Merit-Rating Factor in a Multiplicative Rate-Making Model." *ASTIN Colloquium, Cannes.*

Pinquet, J. (1997). "Allowance for Costs of Claims in Bonus-Malus Systems. " *ASTIN Bulletin, 27, 33-57.*

Pinquet J. (1998). "Allowance for Costs of Claims in Bonus-Malus System. "*ASTIN Bulletin, 27, 33-57.*

Pinquet J. (1998). "Designing Optimal Bonus-Malus Systems From Different Types of Claims". *ASTIN Bulletin, 28, 205-220***.**

Pollard, R., B. Benjamin, and C. Reep. (1977). "Sport and the Negative Binomial Distribution". *In S. Ladany and R. Machol, (eds.), Optimal Strategies in Sports. Amsterdam: North Holland.*

Prins, H.J. and F. Roozenboom (1982). "Bonus/Malus. In New Motor Rating Structure in the Netherlands". *ASTIN Groep Nederland, 72-104.*

Raviv, A. (1979). "The Design of an Optimal Insurance Policy." *American Economic Review. 84-96.*

Reep, C, R. Pollard, and D. Benjamin. (1971). "Skill and Chance in Ball Games." *Journal of the Royal Statistical Society. A, 134, 623-629.*

Renshaw, A.E. (1994). "Modelling The Claims Process in the Presence of Covariates." *ASTIN Bulletin, 24, 265-285.*

Richardson, L. (1945). "The Distribution of Wars in Time." *Journal of the Royal Statistical Society. A, 107, 242-250.*

Ruohonen, M. (1988). "A Model for the Claim Number Process." *ASTIN Bulletin. 18, 57-68.*

Schiesinger, H. (1981). "The Optimal Level of Deducibility in Insurance Contracts."

Journal of Risk and Insurance. 48, 465-481.

Seneta, E. (1981). "A Non-Negative Matrices and Markov Chains". *Springer- Verlag.*

Sharif, A.H., and H.H. Panjer. (1993). "A Probabilistic Model for Automobile Claims: A Comment on the Article by M.N. Islam and P.C. Consul." *Mitteilungen der Vereinigung Sehweizeriseher Versicherungsmathematiker. 279-292.*

Sichel, H. (1971). "On a Family of Discrete Distributions Particularly Suited to Represent Long-Tailed Frequency Data." *Proceedings of the Third Symposium on Mathematical Statistics, N. Loubsher (ed.), Pretoria.*

Sigalotti, L. (1994). "Equilibrium Premiums in a Bonus-Malus System." *ASTIN Colloquium, Cannes.*

Simar, L. (1976). "Maximum Likelihood Estimation of a Compound Poisson Process. " *The Annals of Statistics, 4 1200-1209.*

Stuart, C. (1983). "Pareto-Optimal Deductibles in Property-Liability Insurance: The Case of Homeowner Insurance in Sweden." *Scandinavian Actuarial Journal. 227-238.*

Sundt, B. (1984). "An Introduction to Non-Life Insurance Mathematics." *Karlsruhe: Verlag Versicherungswirtschaft.*

Sundt, B. (1989b). "Bonus Hunger and Credibility Estimators with Geometric Weights." *Insurance: Mathematics and Economics. 8, 119-126.*

Sundt, B. and W. Jewell. (1981). "Further Results on Recursive Evaluation of Compound Distributions." *ASTIN Bulletin. 12, 27-39.*

Taylor, G. (1990). "Setting A Bonus-Malus Scale in the Presence of Other Rating Factors.." *ASTIN Bulletin. 27, 319-327.*

Tremblay, L. (1992). "Using the Poisson Inverse Gaussian in Bonus-Malus Systems." *ASTIN Bulletin. 22. 97-106.*

U.P.E.A. (1992). "La nouvelle tarification en assurance R.C. automobile: Une mise a

Pheure eurooeenne." *Les cahiers de Passurance, Bruxelles.*

van Eeghen, J., J.A. Nijssen, and F.A.M. Ruygt. (1982). "Does a Bonus/Malus System Always Lead to a Premium Crash? A Markovian Analysis." *In New Motor Rating Structure in the Netherlands, ASTIN Groep Nederiand, 120-128.*

van Eeghen, J„ J. Greup, and J.A. Nijssen. (1983). "Surveys of Actuarial Studies: Rate Making." *Research Department, Nationale Nederlanden.*

Vandebroek, M. (1993). "Bonus-Malus System or Partial Coverage to Oppose Moral Hazard Problems?" *Insurance: Mathematics and Economics, 13, 1-5.*

Vandenbroek, M. (1993). "Bonus-malus system or partial coverage to oppose moral hazard problems"*, Insurance: Mathematics and Economics, 13, pp 1-5.*

Venezia, I. and H. Levy (1980). "Optimal Claims in Automobile Insurance." *Review of Economic Studies, 47, 539-549.*

Venezia, I. and H. Levy (1983). "Optimal Multi-Period Insurance Contracts." *Insurance: Mathematics and Economics. 2, 199-208.*

Venter, G. (1990). "Credibility." *In Foundations of Casualty Actuarial Science, Washington, D.C.: Casualty Actuarial Society.*

Venter, G. (1991a). "Effects of Variations from Gamma-Poisson Assumptions." *Proceedings of the Casualty Actuarial Society. 78, 41-55.*

Venter, G. (1991b). "A Comparative Analysis of Most European and Japanese Bonus-malus Systems: Extension." *Journal of Risk and Insurance. 58, 542-547.*

Vepsafainen, S. (1972). "Applications to a Theory of Bonus Systems." *ASTIN Bulletin. 6, 212-221.*

Vrontos S. (1998). "Design of an Optimal Bonus-Malus System in Automobile Insurance." *Msc. Thesis, Department of Statistics, Athens University of Economics and Business, ISBN: 960-7929-21-7.*

Wald, A. and J. Wolfowitz. (1951). "Bayes Solutions of Sequential Decision Problems." *Annals of Mathematical Statistics. 53, 82-99.*

Walhin, J. F. and Paris, J. (1997). "Using mixed Poisson process in connection with bonus-malus systems. " *ASTIN Bulletin, Vol 29, No I 1999, pp. 81-99.*

Walhin, J. F. and Paris, J. (2000). "The true claim amount and frequency distribution within a bonus-malus system. " *ASTIN Bulletin, Vol 30,2000, pp. 391-403.*

Welten, C.P. (1968). "The Unearned No-Claim Bonus." *ASTIN Bulletin.5, 25-32.*

Wiflmot, G.E. and B. Sundt. (1989c). "On Posteriori Probabilities and Moments in Mixed Poisson Processes." *Scandinavian Actuarial Journal, 139-146.*

Wiilmot, G.E. (1990). "AsymptoticTail Behaviour of Poisson Mixtures with Applications." *Advances in Applied Probability. 22, 147-159.*

Wiilmot, G.E. (1993). "On Recursive Evaluation of Mixed Poisson Probabilities and Related Quantities." *Scandinavian Actuarial Journal, 114-133.*

Wiilmot, G.E. and B. Sundt. (1989d). "On Evaluation of the Deiaporte Distribution and Related Distributions." *Scandinavian Actuarial Journal. 101-113.*

Willmot, G.E. (1986). "Mixed Compound Poisson Distributions." *ASTIN Bulletin. 16-S, 59-79.*

Willmot, G.E. (1987). "The Poisson-Inverse Gaussian Distribution as an Alternative to the Negative Binomial." *Scandinavian Actuarial Journal._113-127.*

Willmot, G.E. (1988b). "Sundt and Jewell's Family of Discrete Distributions." *ASTIN Bulletin. 18, 17-29.*