ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

IJAFAME
International Journal of Accounting, Finance, Auditing, Management & Economics

Non-life insurance ratemaking techniques: A literature review of the classic methods

Fatima El Kassimi, ***(Ph.D. student)***
Faculty of Law, Economics and Social Sciences of Settat
Hassan I University of Settat, Morocco

Jamal Zahi, ***(Phd, Professor)***
Head of laboratory of mathematical modeling and economic calculus (LM2CE)
Faculty of Law, Economics and Social Sciences of Settat
Hassan I University of Settat, Morocco

Correspondence address: Laboratory of mathematical modeling and economic calculus LM2CE, UNIVERSITY Hassan the first Faculty of Juridical, Economic and Social Sciences Settat, MOROCCO.
elkassimi.fatima@yahoo.fr

Disclosure statement: The author is not aware of any funding that might be perceived as affecting this study's objectivity.

Conflicts of interest: The author(s) reports no conflicts of interest.

Cite this article El kassimi, F., & Zahi, J. (2021). Non-life insurance ratemaking techniques: A literature review of the classic methods . International Journal of Accounting, Finance, Auditing, Management and Economics, 2(1), 344-361. https://doi.org/10.5281/zenodo.4474479

DOI: 10.5281/zenodo.4474479
Published online: January 29, 2021

Copyright © 2021 – IJAFAME

Non-life insurance ratemaking techniques: A literature review of the classic methods

Abstract:

The primary role of insurance is to protect and guarantee individuals' financial safety and security against the financial consequences of random incidents. It involves aggregating a large number of individual risks, among which there will be a certain amount of insurance claims and accumulated losses to the insurance company during a specific timeframe. Furthermore, one of the insurers' main concerns is establishing a tariff structure that distributes these claims and losses among policyholders most equitably and reasonably. This task of determining the pure premium belongs predominantly to actuaries who evaluate the probability of the risk occurrence, determine the risk factors in order to establish commensurate tariffs for each class so that everyone and each pay premium that, in one way or the other, reflects their riskiness. This article provides an overview of the fundamental concepts of ratemaking and reviews the classical statistical techniques used by insurance companies to discriminate tariffs. The article follows the customary actuarial distinction between the two main pricing techniques, namely the a priori and the a posteriori ratemaking techniques.

Keywords: Actuarial sciences, Non-Life Insurance, Ratemaking, Tariff Class; Risk Classification.
JEL Classification : (G22)
Paper type: Theoretical Research.

ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

IJAFAME
International Journal of Accounting, Finance, Auditing, Management & Economics

1. Introduction

The founding idea behind insurance is the protection of individuals against the financial consequences of damaging random events. Through sharing danger or risk among a group of individuals, the damage related to this danger is most often so significant that an individual would not be able to bear it on his own (the risk of being affected by a chronic disease or the risk of accident for example). On the other hand, if this danger is shared among the individuals, the loss becomes bearable for each one. This insurance principle requires grouping individual risks into various categories or classes with a homogeneous set of characteristics. Such classification usually boils down to different classes for which members of a given class share the same set of risk characteristics; in that respect, each class will have a certain number of insurance claims and accumulated losses and pay the same premium rate. To reasonably estimate this premium and price insurance policies, insurers must predict the expected loss accurately, often referred to as the "pure premium."

In reality, despite using classification methods, the risk is rarely distributed homogeneously throughout the group. Furthermore, Insurers tend to select individuals with "good risks" from the group. If the selection is sufficient, the insurance company will have fewer claims to pay and can, therefore, offer a better rate to its policyholders. On the other hand, the rest of the insurers are left with the "bad risks," which leads them to increase their rates. One way to avoid this split in the insurance market is that the rest of the insurers have to follow the first insurer; otherwise, they will be forced to offer different rates between good and bad risks to keep their clients. The insurance market, therefore, tends to differentiate rates according to the degree of risk. The inherent trend in insurance obliges actuaries to invent and develop somewhat effective differentiation methods for calculating rates.

One significant aspect of the insurance business is the estimation of the pure premium, which can be defined as the tariff to pay in exchange for the risk transfer. To estimate such premium, Actuaries call upon probability and statistics to accurately measure and manage risks related to the occurrence of events. One customary practice in all insurance areas could be observing past events forecast future claims and costs. For instance, actuaries classify insured based on individual characteristics; the estimate differentiated premiums within the insurance portfolio. This step is referred to as a priori analysis, which involves the classification of the risk ensembles provided the influencing factors so that each class's insureds pay the same premium rate.

Moreover, this step is completed by estimating the pure premium, which is mathematically formulated as the product of the claims frequency's conditional expectation and the expected cost of claims. The second step is known as *a posteriori* analysis. As its name suggests, historical data of the insured's past claims is used. Making use of such information allows for correcting and adjusting the *a priori* tariff calculated beforehand.

The present article aims to present a literature review of the essential classical ratemaking models and discuss the issues related to establishing prices in insurance markets. Providing the fore-mentioned purpose, the article is organized as follows: the second section presents the concept of ratemaking in non-life insurance, highlighting the information issues within the insurance industry, namely: moral hazard and adverse selection. Furthermore, it presents the unique techniques that the insurance companies use to counterparts such problems, mainly risk classification and differentiated tariff structures. Section 3 highlights the difference between a priori and a posteriori ratemaking methods giving a review of the various empirical studies, reviewing the core classical statistical techniques of pricing risks in non-life insurance. In Section 4, we conclude with some remarks.

2. Ratemaking concept and information issues :

The premium, rate, or tariff is the price charged for the risk transfer from the insured to the insurance company. Essentially, it has a double function. First, it should produce total funds sufficient to cover the insurer's obligation; second, it should fairly share the insurance cost among insureds. We usually expect that a rate will meet both requirements, but the first function is undoubtedly far more critical, as the primary obligation of an insurer is to remain solvent at any given time. Actuarially speaking, this translates to the average premium being adequate, meaning at least sufficient to meet losses, including administrative expenses. The second primary requirement of a sound rate is that it should distribute insurance costs equitably among insureds. Another way of saying this is that the premium should represent a fair allocation of costs among insured risks and therefore be equitable. In this respect, we can think of insurance as a device to pool risks and share losses. Consequently, the ratemaking process can be thought of as an ensemble of techniques used by the insurer to estimate the price paid by the insured in exchange for transferring their risk.

Within a heterogeneous insurance portfolio, not all insureds are equal regarding their risks; some present a more dangerous profile than others. Therefore, charging the same premium for all might seem unfair, as it would necessarily lead to some insureds being overpriced and using these premium surcharges to compensate for claims caused by riskier individuals. The heterogeneity of the portfolio leads directly to the so-called anti-selection phenomenon, also known as adverse selection, which is a form of asymmetric information. This particular type of information issue is expected in the insurance markets. Typically, the insureds have a better knowledge of their characteristics and behaviors than their insurer does. The two well-known information problems discussed in the economics literature are adverse selection and moral hazard (Arrow, 1963). The concept of anti-selection refers to situations where, prior to the conclusion of an insurance contract, the policyholders have an information advantage over the insurer. Denuit (2007) considers that clients have a better knowledge of their claim behavior than insurance companies do; the insurer may use deductibles to separate individuals with different risk levels. On the other hand, Chiappori, Jullien, Salanié, and Salanié (2006) stressed that moral hazard occur when the outcome of a claim occurrence or an accident depends, in an unexpected way, on a decision that is in most cases made by one party behind closed doors and thus not observable by the other. Classically, the insurance company party usually chooses to make an expensive effort to diminish its risk.

2.1 Information issues and risk classification:

In the literature on insurance, empirical studies have mainly focused on the positive correlation between risk, the accident probability of a policyholder, and the level of insurance coverage. As well documented in Chiappori (2000), moral hazard and adverse selection should introduce a positive correlation between accidents and insurance coverage. Presuming charging different tariffs to identical agents, it follows that the frequency of accidents among the subscribers of a given contract would increase following the coverage level. As predicted by the Rothshild and Stiglitz (1976) model of competition under adverse selection and Wilson (1977) model, this correlation stems from the fact that "high-risk" agents are usually willing to pay more for extra coverage than those with "low-risk," and will consequently choose contracts with higher coverage leading to accidents. Shavell (1979), Holmstrom (1979), and Arnott and Stiglitz (1988) made the same conclusion regarding this positive correlation but this time under a moral hazard; where an opposite causality leads to the same correlation: an insured who, for any unspecified reason, switches to a high coverage contract have a weaker incentive for safe driving and therefore becomes riskier. Within the automobile insurance markets. Puelz and Snow (1994) confirmed, as predicted in theory, the presence of a positive link between the claims' number and coverage contracted within a risk class in the portfolio of an American

ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

IJAFAME
International Journal of Accounting, Finance, Auditing, Management & Economics

insurer. However, in recent studies, Chiappori and Salanié (2000) and Dionne et al. (2001) found no evidence of asymmetric information among French and Canadian insurers, respectively.

According to Dionne, Michaud, and Pinquet (2013), the main difference between the two phenomena is that anti-selection effects, in one way or the other, the insurance contracts. In contrast, moral hazard affects individuals' hidden actions. In this respect, the information problems in the insurance market can be regarded as the effect of applying the same tariff for the entire heterogeneous portfolio, which implies ensuring high-risk contract at a lower price (that does not reflect their true riskiness) compared to their real cost and consequently discouraging insuring medium risks.

The presence of the moral hazard and adverse selection in the insurance market could lead to situations where the insurance company could end up with a large number of high risks in its portfolio, which may subsequently lead to continually increasing tariff rates to keep the company solvency; one cannot stress enough the importance of pricing for non-life insurance. Economic theory, mainly Chiappori, Salanié, and Dionne (2000), Gouriéroux and Vanasse (2001, 2006), teaches us that to reduce the risk of adverse selection, the insurer should divide the insurance portfolio into sub-portfolios based on certain influence factors. Therefore, every class will contain policyholders with an identical risk profile that will automatically pay the same tariff. The risk classification technique is based on this principle; it allows for reasonably pricing products using a credible statistical basis.

2.2 Risk classification and tariff structure:

In constructing a risk classification system, Actuaries usually use econometric models, mainly linear regression (linear regression is used to estimate the effect of explanatory variables over the phenomenon of interest) to establish a tariff structure that reflects the various risk profiles in a portfolio. Such models make it easier for the actuary to include various classifying variables, to create risk classes corresponding to each risk profile. When selecting explanatory variables, statistical considerations are of great importance and must be considered while choosing the classifying variables. Thus, every explanatory variable should meet the following three actuarial criteria: first, be accurate, meaning that it has a direct impact on costs; second, meet homogeneity prerequisite so that within a given class, the expected loss is more or less similar, and finally yet importantly be statistically sound and reliable. See Finger (2001) for more information on these criteria.

Typical response variables are claim frequency per unit of exposure for property-casualty insurance and the corresponding claim severity. Various classifying variables can be found in all shapes of insurance. For instance, in the car insurance case, it is customary for actuaries to cross-classify risks based on various factors: Drivers are often categorized according to a driver-class variable formed from a composite of individual characteristics, including age, gender, occupation, prior accident experience. Vehicles are classified in terms of age, model, value.

The purpose of classification is to allow for the most accurate prediction of each individual's expected pure premiums by eliminating cross-subsidy between insured with low and high risks. However, from the 70s on issues related to estimating pure premiums using one or more dimensions have been primarily debated [e.g., Bailey (1963) Bailey and Simon (1960) Chang and Fairley (1979), Sant (1980), Weisberg and Tomberlin (1982), Weisberg, Tomberlin, and Chatterjee (1984)]. Moreover, since the information used to estimate premiums in the future usually consists of, for a given population and over a particular period, policyholders' claim experience, the adjustments based on multiple classifications could be very sensitive to extreme values in case some class had only limited exposures.

3. Non-life insurance rate making techniques:

Historically, back in the 19th century, standard Gaussian linear regression was a custom practice among actuaries. The linear econometric model, suggested by Legendre and Gauss, made it possible to measure the exogenous variables' impact over the studied phenomenon.

However, linear modeling involves a series of hypotheses in order for it to function; three out of seven hypothesis, namely the normality, the homoscedasticity, and the Gaussian probability density hypothesis, are far from being valid, making it hard to apply the model, due to the stochastic nature of risk occurrence in insurance that directly affects the frequency–severity distribution.

3.1 A priori ratemaking:

A priori pricing is a pricing method used by actuaries to better segment insurance portfolios. This method predicts the expected number of claims based on the insured's observable characteristics such as age, gender, mileage, vehicle use, and occupancy; the expected number of claims is predicted based on the insureds' observable characteristics.

In 1960, Bailey and Simon believed that the purpose of a priori pricing is to construct homogeneous risk classes where insured belonging to the same risk class pay the same premium. We can define risk classes as a set of characteristics, where insured belonging to the same risk class have identical observable characteristics. These observable characteristics of insureds are called classification variables or a priori variables. In general, a priori classification is done using regression models or general linear models (GLM) developed by (Nelder and Wedderburn, 1972).

The purpose of regression is to analyze the relationship between the response variable and the explanatory variables. For instance, in automobile insurance, the response variable generally represents the number of claims or claims costs, while the explanatory variables represent the classification variables. We can express this relationship by an equation that predicts the response variable using a function (linear combination) involving the explanatory variables and the prediction parameters. In a priori pricing, the number of claims is generally modeled via Poisson, negative binomial distributions, while the cost of claims is generally modeled via gamma, inverse-Gaussian distributions. Risk classification techniques have been the subject of several articles in the actuarial literature. Among others (Dionne and Vanasse, 1989) used a negative binomial regression model, and (Dean, Lawless, and Willmot,1989) used the Poisson-inverse Gaussian distribution to model the number of claims.

3.1.1 Generalized linear models:

In the 1960s, Canadian actuaries developed a pricing method known as the Minimum bias procedure (Bailey and Simon, 1960). The principle of this method is to define a relationship between the explanatory variables, risk classes, and the distance between predicted and observed values. After setting these elements, we can then calculate the coefficient to be associated with each risk level via an iterative algorithm using minimizing distance. Later these algorithms proved to be exceptional cases of Linear Generalized Models.

For a long time, actuaries limited themselves to using the Gaussian linear model when quantifying the impact of explanatory variables on a phenomenon of interest (frequency or cost of claims, probability of occurrence of insured events). Now that the complexity of the statistical problems facing actuaries has increased considerably, it is crucial to turn to models that take better account of insurance's reality than the linear model does. Indeed, the latter imposes a series of limitations" assumptions" that are difficult to reconcile with the reality of the numbers or cost of claims: (approximately) Gaussian probability density, the score's linearity, and homoscedasticity. Even if it is possible to overcome some of these constraints by transforming the response variable beforehand using well-chosen functions, the linear approach

ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

has many disadvantages (working on an artificial scale and difficulties returning to the initial quantities).

A first step in using models more appropriate to insurance reality was taken when the London actuaries of City University applied Generalised Linear Models (GLM) in actuarial sciences at the end of the 20th century. These models, introduced into statistics and actuarial sciences by John Nelder and Robert Wedderburn (1972), make it possible to avoid the hypothesis of normality by extending the Gaussian model to a particular family of distribution, specifically the exponential family (which, in addition to the customary law, includes the Poisson, binomial, Gamma and Inverse Gaussian laws). For further discussion, see Gouriéroux, Monfort, and Trognon (1984).

Poisson's regression (and apparent models, such as the negative binomial regression) is now a tool of choice for automobile pricing, mostly succeeding the general linear model and logistic regression in analyzing claim counts. The method had its breakthrough following its inclusion in the most commonly used statistical software (SAS in particular). Subsequently, the GLMs have been widely used in non-life insurance and have become a standard industry practice for pricing risks. They are now commonly used for estimating the pure premium through the frequency–severity approach, based on a priori characteristics of the insurance policy.

In addition to the maximum likelihood approach, GLM techniques allow the analysis of several phenomena from a quasi-likelihood perspective by specifying only the mean-variance structure. In this respect, French econometricians have proved fundamental results of the convergence of estimators obtained in this way. See, in particular, Gouriéroux, Monfort, and Trognon (1984).

Nelder continues developing the GLMs theory together with Peter McCullagh (1989), together they worked on several regression models used in actuarial sciences. Moreover, GLMs have undergone a succession of important papers; for further discussion, refer to Denuit et al. (2007), Antonio and Beirlant (2007), Ohlsson (2008), Jong and Heller (2008), Frees (2009), Ohlsson, and Johansson (2010). Nevertheless, we can never have a full opinion about a model until we have used it enough to prove its strength and weaknesses, GLMs have been the subject of study among scholars throughout the years, and as any other model, they proved their shortcomings, specifically when it comes to their application. Pregibon et al. (1984) proposed a dual generalized linear model (DGLM); meanwhile, the maximum likelihood estimation of DGLM was later applied to non-life insurance pricing was introduced initially by Smyth. Aitkin et al. (1989) present a proper application of the GLMs in estimating premiums; they study several application examples of GLMs, including the Poisson distribution, Ohlsson and Johansson (2010) introduced the generalized linear model with a practical application in automobile insurance in which the claim frequency is fitted with a Poisson distribution model, while gamma distribution model is used for fitting claim severity.

GLMs extend the framework of ordinary (standard) linear models to the distributions class derived from the exponential, making it possible to model different measures such as binary, skewed data, and counts. Should the reader need further information, Haberman and Renshaw (1996) give an overview of GLMs' applications in actuarial sciences. Further discussion can be found in Kaas et al. (2008), Jong and Heller (2008), and Frees (2010). The initial GLM models introduced in the statistics literature find their origins in Nelder and Wedderburn (1972).

The following table summarizes some of the central studies using the Generalized linear models and some of the models' most recent usage.

Table1. Studies featuring Generalized linear models

Year	Authors	Article	Main ideas	Model
1960	*-Robert Bailey* *-LeRoy Simon*	"Two Studies In Automobile Insurance."	They examined models with multiplicative and additive functions via balance principle and Âsquared bias functions, using loss costs and loss ratios. They used the Minimum bias procedure to compare additive and multiplicative classification models for the *Canadian private passenger automobile business*. They discuss the rationale for the minimum bias procedure, the characteristics of a suitable rating model, and the rating scenarios that fit various models.	-Minimum bias procedure
1972	*-John Nelder* *-Robert Wedderburn*		The authors show that The maximum likelihood estimates for a large class of commonly used regression models can be obtained via Iteratively weighted least squares, in which both the weights and the response are adjusted from one iteration to the next. The proposed algorithm, also known as "fisher score."	-Generalized Linear Models
1989	*McCullagh, P. and Nelder, J.A.*	Generalized Linear Models	The paper presents an approach to estimate the dispersion parameters for bivariate, trivariate, and multivariate correlated binary data, with scalar and matrix values. -they present numerous studies showcasing the impact of over-dispersion on the univariate data analysis, comparing other approaches with these studies. -They used these estimates to adjust the correlated binary data, using *the Hunua Ranges data as an ecology problem.*	-Generalized Linear Models
2007	*-Michel Denuit* *-Xavier Maréchal* *-Sandra Pitrebois* *-Jean-François Walhin*	Actuarial Modelling Of Claim Counts Risk Classification, Credibility And Bonus-Malus Systems	The book is devoted to analyzing the number of claims filed by an insured driver over time. - The authors make extensive use of the generalized linear models, along with the bonus-malus system in pricing claims.	-Generalized Linear Models
2008	*Piet de Jong* *Gillian Z. Heller*	Generalized Linear Models For Insurance Data	The book introduces the GLMs to support critical decisions within the insurance industry And addresses the complications that insurance data entails. The book covers all customary exponential family distributions, extends the practice to correlated data structures, and presents contemporary developments. Selection in the presence of enormous data sets and the treatment of fluctuating exposure times are also discussed in the book. -It also provides practical data-based Exercises to help readers consolidate their skills	- Generalized Linear Models and their extensions.

ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

2010	*-Ohlsson, Esbjörn,* *-Johansson, Björn*	Non-Life Insurance Pricing With Generalized Linear Models	The book focuses on the basic theory of generalized linear models (GLMs) in a pricing analysis setting, giving some useful extensions that are not commonly used. The book provides countless theoretical results accompanied by numerous examples and mathematical illustrations. The textbook is used in actuarial education and is intended for practicing actuaries with a solid mathematical and statistical background."	-Generalized Linear Models and their extensions.
2016	*-Filipe Charters de Azevedo* *-Teresa A. Oliveira* *-Amilcar Oliveira*	Modeling Non-Life Insurance Price For Risk Without Historical Information	The article generally presents the ratemaking mechanisms in non-life insurance through the GLM regression models, specifically the Gamma, Poisson, and Tweedie models. Given the difficulty of applying these models in experimental design, they used Box-Cox transformation with Seemingly Unrelated Regression. An application of these techniques in the motor is presented.	-Generalized Linear Models, - Poisson, Gamma, and Tweedie models.
2017	*-Julien Antunes Mendes,* *-Sébastien de Valeriola,* *- Samuel Mahy* *- Xavier Maréchal*	Machine Learning applications Frequency modelling: An educational case study	This paper presents a comparison on a simulated database between traditional statistical predictive modeling techniques (Generalized Linear Models and generalized additive models), machine learning techniques (regression trees, bagging, random forests, boosting, and neural networks), and penalized regression techniques (Lasso, Ridge and Elastic Net).	- Generalized Linear Models -Generalized Additive Models -Machine Learning
2018	*-Giorgio Alfredo Spedicato,* *-Christophe Dutang,* *-Leonardo Petrini*	Machine Learning Methods to Perform Pricing Optimization. A Comparison with Standard GLMs	This article explores the applicability of the novel machine learning techniques such as tree-boosted models in optimizing premiums on prospective policyholders, comparing the results with the classic Generalized Linear Models GLMs.	-Machine Learning -Generalized Linear Models
2019	*- J. Zhang,* *-T. Miljković*	Ratemaking for a New Territory: Enhancing GLM Pricing Model with a Bayesian Analysis	-This paper provides a Bayesian approach in ratemaking for companies desiring to start a new business in a new territory or a somewhat new territory to have partial claims experience. -A Bayesian Poisson regression model is suggested for modeling claims frequency. Bayesian analysis of claim severity considers a gamma regression and non-informative uniform priors for the regression coefficients. - Conclusions. Bayesian analysis with power prior can effectively use auto insurance ratemaking for promoting new business in a new territory or improve a growing business's pricing in a new territory.	-Generalized Linear Models with a Bayesian Analysis.

2020	*-Mario V. Wuthrich* *-Christoph Buser*	Data Analytics for Non-Life Insurance Pricing	This paper aims at giving a comprehensive skill set to actuaries in insurance pricing and data science. It starts from the classical basic generalized linear models, generalized additive models, and credibility theory. These methods are the building blocks of the deeper statistical understanding. Then it presents several machine-learning techniques such as regression trees, bagging, random forest, boosting machines, and neural networks.	- Generalized Linear Models -Generalized Additive Models - Credibility Theory -Machine Learning -Data Analytics

Source: *created by the authors*

3.2 A posteriori rating:

During the construction of an a priori tariff structure, if all the characteristics influencing the number of claims could be measurable and incorporated into pricing, the risk classes would likely be homogeneous. Individual differences from the average would not cause a premium adjustment. Unfortunately, many of the insured characteristics cannot be considered in a priori pricing either because they are not observable or challenging to measure. These include, among others, reflexes, driving under the influence of aggressive alcohol driving; for instance, the aggressiveness of a driver (insured) behind the wheel is difficult to assess. However, as the driver's claims history starts taking shape, more information becomes available, and the true riskiness of the policyholder becomes more visible and quantifiable to the insurer. The optimal predictions can be constructed from both the initial (or a priori) risk classification and the past claims' summary statistics. In this respect, the premium will be regularly updated as more data are collected in the future. It is well known that these hidden variables can significantly impact the number of claims by policyholders Denuit et al. (2007). As a result, the portfolio is still heterogeneous despite using several classification variables in a priori pricing.

Thus, actuaries use a rating method based on the insured's claims experience. Known as an ex-post rating, actuaries use it to take into account the individual differences of each insured in the portfolio. This method consists of modeling the heterogeneity of the portfolio using a random effect. A posteriori analysis of this random effect as a function of the insured's number of claims allows the premium to be re-evaluated a priori to reflect the real risk represented by the insured. The use of claims history to adjust the insured's premium comes from the fact that the best predictor of the number of future accidents the insured will report is not the age or the type of vehicle but the number of past accidents reported by the insured Denuit et al. (2007).

3.2.2 Credibility theory:

In actuarial science, the advent of credibility dates back to Arthur. H Mowbray (1914) and A.Whitney (1918). These two the first to apply the credibility theory for pricing purposes. Their work gave rise to the theory of limited fluctuations, also known as the American credibility, which is based on suggestions to answer the question raised in the General Motors affair with its insurer Allstate(1910)1: from-what size can a company be priced exclusively based on its own experience? Mowbray (1914) was the first to provide a clear answer to this question, laying down the foundations for the so-called theory of Limited fluctuations. However, he only proposed a sufficient threshold from which a company is considered large enough, without mentioning what would happen to smaller firms; a few years later; A.Whitney (1918) states "the necessity, for equity purposes, to weigh collective experience, on the one hand, and

[1] Back in 1910, *Allstate* insured General Motors and a number of small businesses against work accidents. By calculating the average premium rate based on its experience, General Motors realized that its premium should be lower than that of all insured firms. Claiming that the number of insured was large enough, subsequently it required its insurer to take into account its own history rather than that of all insured.

ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

IJAFAME
International Journal of Accounting, Finance, Auditing, Management & Economics

individual experience on the other." The whole theory of credibility aims precisely at calculating this weighting in the best possible way.

According to Waters (1994), Credibility theory is a set of techniques used to calculate premiums for short-term contracts, utilizing two elements: past data from the risk itself and collateral data, i.e., data from other relevant sources. The primary purpose of credibility models is to predict the future loss for a particular policy, using previously observed data. The credibility premium formula as derived by Waters (1987) is of the form; m = ZX + (1 - Z) Where; m is the premium, z is the weight or credibility factor, the mean parameter x is the observed mean claim amounts per unit risk exposed for individual risk itself. The credibility factor Z is a measure of how much reliance the company is willing to place on the data from the policy itself. It ranges from 0 to 1. In credibility of data, 0 credibility is given to data that is too small to be used for premium rate making. If some data has a credibility of 1, this means the data is entirely credible. In this sense, Z value reflects how much "trust" is placed in the data from the risk itself compared with the larger group's data.

North American actuaries originally developed credibility theory in the middle of the 20th century. An early beginning of the theory appeared in Mowbray (1914) as a premium calculation technique. It assumes that the yearly claims are independently and identically distributed random variables from a probabilistic model with defined means and variance, assuming that the data follows a normal distribution. Whitney (1918) and other researchers criticized this theory. Whitney suggested that claims are random in nature, and hence the assumption of the fixed effects model was invalid. The theory also faced partial credibility since it was difficult to determine the credibility factor's value.

After World War II revolution, Whitney's random effect model came into place. Later on, Nelder and Verall (1997) derived credibility functions by the generalized linear model approach and consequently included the random-effects model. This has provided a wide range of actuarial applications; among them are premium rating and reserving. Though much research was done that yield several findings, it was found that the fixed effect credibility could not solve the problem of credibility. It is said that part of it was due to undeveloped or poor statistical background.

The Greatest accuracy credibility theory method, also called the Least Squares, or Bühlman's Credibility originates from Bailey's two seminal papers (1945,1950); it uses both the variance of observations within each company and the variance across one company to another. In his 1945 paper, Bailey obtains a credibility formula that seems to anticipate the non-parametric universe to be explored two decades later by Bühlman. The application of such technique typically requires estimating the so-called Bühlman's Credibility Parameter K. However, this technique is theoretically complete and meets the criteria for a credibility method, however, there is one shortcoming to the industry experience it can be challenging to acquire additional information.

Bayesian ideas and techniques were introduced into actuarial science in a big way in the late 1960s when Bühlman (1967, 1969) papers and Bühlman Straub (1970) laid down the foundation to the Bayes credibility approach. The Bühlman credibility model uses a Bayesian framework and assumes each unit's risk parameters to be independent and follow a normal distribution. Besides, conditional on a unit's risk parameter, the losses are considered independent and identically distributed. The expected quadratic loss of a linear predictor is then minimized to produce the Bühlman credibility premium. However, a great deal of work has been done to extend Bühlman's model, including Bühlman Straub (1970), who generalize the model in cases where volume is involved, and the Hachemeister regression credibility model Hachemeister (1975), which introduces covariates to the conditional mean of losses. Klugman et al.(2008) Denuit et al. (2007) studied the frequency distribution of insurance claims and parameter estimation methods. Bühlman (1967) presented the credibility approach in the form

of a linear function to estimate and predict future periods' expected claim counts, using past data on claims as a risk class or connected risk classes. Bühlman's credibility theory is fascinating and can be extended to other approaches, such as the Bühlman-Straub model, Jewell's model, or the exact credibility approach.

Although the credibility theory combines different data collections to provide an accurate overall estimate, it is problematic to implement due to its mathematical complexity Denuit (2006). As a result, simplified versions of credibility theory, namely the bonus-malus system introduced by Pesonen (1962), where the surcharges for reported claims are referred to as maluses and bonuses being the discounts for claim-free periods. The bonus-malus system is sometimes used as part of an overall strategy aimed at retaining profitable customers, as has been evidenced by Pitrebois et al. (2003), where "good" customers can be defined as being equivalent to profitable customers, i.e., those for whom observed losses over the years are lower than expected. According to Denuit (2006), these systems allow premiums to be adapted for hidden individual risk factors through the history of past claims. Therefore, in the context of insurance markets, the bonus-malus system's primary purpose is to assess the individual degree of risk equitably so that the insurance company will demand a premium corresponding to the insured risk profile and claim history.

3.2.2 Bonus- Malus system

Actuaries use several mathematical systems or concepts to estimate the premium for insureds based on their claims experience. These experience-rating systems penalize insureds responsible for one or more reported accidents with surcharges (maluses) and rewarded insured with no claims with discounts (bonuses). From an insured's perspective, it is not always clear how the insurer determines discounts and surcharges based on their claims history. For this reason, actuaries have developed a new method of retrospective pricing known as the bonus-malus system. The purpose of this system is to determine, in an appropriate manner but also in a way that is understandable to the broad public (insured, agents, brokers, directors, officers), the amount of premium to be allocated to each insured based on their claims history. The first bonus-malus systems were used in automobile insurance and date back to 1910 in England, followed closely by Canada in 1930, Lemaire (1995). These systems granted a 10% discount, for example, for a year spent without a claim. In case a claim was reported, then no penalty was applied. Since then, bonus-malus systems have evolved considerably, and a theory based on Markov chains has made it possible to analyze them better. Their main advantage is that they offer a simple way to account for a posteriori pricing variables while rewarding policyholders who drive safely. Bonus-malus systems are mainly used in automobile insurance since it is generally accepted that drivers have some control over their accident rate. It is in automobile insurance that this theory has most developed and has acquired its terminology. This principle is also used in reinsurance and group insurance.

Throughout the world, either the government imposes bonus-malus systems, or the market is completely free. When they are imposed, all insurers must adopt the same system. On the contrary, when the market is completely free, each insurer builds their private system. In Europe, a free-market law is being implemented, while in Asian countries, bonus-malus are generally regulated by the government, Lemaire (2004). In America, both types are found. In Quebec's particular case, the SAAQI uses a system similar to the bonus-malus to penalize traffic violations. The configuration of the systems also varies around the world. Some are very simple and only consider the number of claims, while others also consider the severity of accidents, the possibility of no increase in the premium, and free coverage Lemaire (1995, 2004).

In a study conducted by Lemaire (1977), he demonstrated that a posteriori variables such as the number of claims, the number of accidents, or the number of traffic violations are much better

ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

predictors of risks than a priori variables. For this reason, it is crucial to incorporate a posteriori variables in the pricing system.
The following table summarizes some of the central studies using the credibility theory and the bonus-malus system and some of the models' most recent usage.

Table 2. Studies featuring Credibility theory and Bonus-Malus system

Year	Authors	Article	Main ideas	Model
1914	*Mowbray*	How Extensive A Payroll Exposure Is Necessary To Give A Dependable Pure Premium. Proceedings Of The Casualty Actuarial Society	In his original paper on Credibility theory, Mowbray suggests how to determine the amount of individual risk exposure needed for $\hat{m}$ to be a thoroughly reliable estimate of m. The problem may be expressed as follows. Suppose that an insured has incurred X_j claims or losses in a given period j where $j\epsilon\{1,2,3,\ldots,n\}$ Assuming that the premium to charge is E(X_j) = ξ , and Var(X_j) = σ^2. the average would be $\bar{X} = n - 1(X_1 + \cdots + X_n)$.We know that E ($\bar{X}$) = ξ, and if the X_j are independent, Var ($\bar{X}$) = σ2/n. The insurer's purpose is to determine the value of ξ.	-The limited fluctuation approach.
1997	*- J A Nelder* *- R J Verrall*	Credibility Theory And Generalized Linear Models	They incorporated the credibility theory within the Hierarchical Generalized Linear Models. Thanks to the framework of the GLMs in general and the Hierarchical Generalized Linear Models in particular, they allow the use of wide-ranging models, they included the random effects in the model to obtain credibility estimates. the paper contributes a further variety of models that may be useful in several actuarial applications, including premium rating and claims reserving	- Credibility theory - Hierarchical Generalized Linear Models
2006	*Ragnar Norberg,*	Credibility Theory	The article reviews the credibility theory, its genesis and origins ; it gives a minute description of its history and development, examining a plethora of studies and extensions of the theory that contributed to its development, revealing its key concerns and results.	-Credibility Theory
2009	*E. Gomez-Déniz*	Some Bayesian Credibility Premiums Obtained By Using Posterior Regret ¡-Minimax Methodology	E. Gomez-déniz makes use of the Bayesian theory to present an application of a unique procedure constructed based on the posterior regret ¡-minimax principle to directly derive a new credibility method via simple classes of distributions. This procedure is applied to the most commonly used pricing principles in insurance, viz. the net, Esscher, and variance principles.	-Bayesian Credibility theory
2010	*Amir T. Payandeh Najafabadi*	A New Approach To The Credibility Formula	This article makes use of the mean square error minimization technique to create a simple yet practical method to the credibility theory. Viz., the Bayes estimator concerning a general loss function and general prior distribution by a convex combination of the observation mean and mean of approximate credibility method. Modification of the approximate credibility for various situations and its form for numerous consequential losses are given.	-Mean square error minimization technique. -approximate credibility method

2010	*-Bohdan Linda, -Jana Kubanová*	Credibility Premium Calculation In Motor Third-Party Liability Insurance	The Bühlmann-Straub credibility model offers a plethora of possibilities. The paper presents several Bühlmann-Straub model applications using various motor third-party liability insurance datasets belonging to five different Slovak insurance companies.	- the Bühlmann-Straub credibility model
2016	*-Serpil Ergün Bülbül - Kemal B. Baykal*	Optimal Bonus-Malus System Design in Motor Third-Party Liability Insurance in Turkey: Negative Binomial Model	Considering the shortcomings of using a mandatory bonus-malus system in turkey, The study uses experience rating through the insured's individual claim experience. Optimal bonus-malus tariffs were estimated via the negative binomial model, credibility theory, Bayesian approach, and the expected value premium principle. The study Uses Claim frequencies data for motor third-party liability policies.	- bonus-malus system -negative binomial model, -credibility theory -Bayesian approach
2019	*-Anani Lotsi -Felix Okoe Mettle - Paul Kwame Adjorlolo*	Application Of Bühlmanns-Straub Credibility Theory In Determining The Effect Of Frequency-Severity On Credibility Premium Estimation	The study uses secondary data of non-life marine insurers in Ghana; it uses the Bühlmans-Straub model to estimate the credibility frequency-severity claim cost. The study compares the resulting premiums and found that the credibility claim costs underestimate claim costs as opposed to the credibility frequency-severity claim costs for the majority of the risk classes. The study Stresses that the inconsistency of claim frequencies and severities with different risk profiles undermine credibility claim costs. The study recommends that credibility pricing based on inadequate claim history or with class risk variation must incorporate credibility risk frequency and severity to determine credibility risk premiums.	-Bühlmans-Straub -Credibility theory
2020	*Olivier Le Courtois*	Q-Credibility	The article adjusts the credibility theory using quadratic adjustment that considers the past observations' squared values. This method introduces non-linearities in the framework. It first describes the overall parametric approach, examining the Poisson-gamma and Poisson-single Pareto distributions. It estimates premiums based only on data, without fitting any distribution, according to the non-parametric approach. It also examines the semi-parametric method where Poisson is the conditional distribution with an unidentified unconditional distribution.	-Quadratic credibility theory

***Source**: created by the authors*

4. Conclusions

The purpose of this article is to give an insight on non-life insurance ratemaking techniques and their development throughout the years, providing the peculiarity of this research field a walk through the basic concepts was necessary to have a better understanding of the theoretical frameworks.

Throughout this paper, we stressed the importance of risk classification in establishing a fair and reasonable tariff structure; in fact, within a heterogeneous insurance portfolio, not all insureds are equal when it comes to their riskiness; some are more risky profile than others are.

ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

IJAFAME
International Journal of Accounting, Finance, Auditing, Management & Economics

Therefore, charging the same premium to all might seem unfair. This heterogeneity can, of course, be reduced by using risk classes that are as homogeneous as possible (based on sex, age, and other characteristics). Given this risk classification, the pure premium for each risk class is estimated using a priori techniques.

The present article provides a literature review on the different pricing methods used in non-life insurance, namely the *a priori* and *a posteriori* classification methods. In the *a priori pricing*, the insurer determines the premium based on the insurer characteristics (such as sex and age), a little is known about the insured, in terms of driving habits, and other behavioral characteristics (non-observable factors), Generalized Linear Models are standard techniques used for a priori pricing. Since observable factors are far from thoroughly explaining the insured's dangerousness, it is therefore quite natural to use an individual's relative claims experience to reassess the amount of their premium. Such is the idea behind *a posteriori* pricing. The insurer uses historical data to correct and adjust these *a priori* premiums using the credibility theory, mainly the bonus-malus system; the premium can then be calculated by multiplying the estimated risk for the following period by the bonus-malus coefficient calculated earlier. In that sense, the criteria of *a posteriori* technique change the risk perception and therefore encourage the policyholders to adopt a more cautious behavior.

The empirical actuarial literature demonstrates the importance of a healthy pricing process for insurance companies. However, this process is slowly getting more challenging, nearly owing to the big-data transformation that the insurance industry is currently undergoing. We can see how data science shapes the insurance industry's future and how these skills are becoming more and more essential for actuaries to acquire. The use of Analytics and Machine learning techniques for pricing purposes is slowly becoming customary practice among actuaries.

References :

(1) A. Lotsi, F. O. Mettle, and P. K. Adjorlolo, "Application of Bühlmanns-Straub Credibility Theory in Determining the Effect of Frequency-Severity on Credibility Premium Estimation," ADRRIJPNS, vol. 3, no. 2, pp. 1-24, Dec. 2019.
(2) Antoine Paglia, Martial Phélippé-Guinvarc'h Tarification Des Risques En Assurance Non-Vie, Une Approche Par Modèle D'apprentissage Statistique.
(3) Antoine Paglia, Martial Phélippé-Guinvarc'h, Philippe Lenca .Adaptation de l'algorithme CART pour la tarification des risques en assurance non-vie.
(4) Antonio, K., Valdez, E.A., 2012. Statistical Concepts of A Priori and A Posteriori Risk Classification in Insurance. Advances in Statistical Analysis 96(2), 187-224.
(5) B. Linda, J. Kubanová, Credibility Premium Calculation in Motor Third-Party Liability Insurance, WSEAS conference Advances in Mathematical and Computational Methods, Malta, 2012, pp. 259-263.
(6) Bailey, R. A., 1963. Insurance Rates with Minimum Bias. Proceedings of the Casualty Actuarial Society 50, 4-14.
(7) Bailey, R.A., Simon, L.R.J., 1960. Two Studies in Automobile Insurance Ratemaking. ASTIN Bulletin 1(4), 192-217.
(8) Bailey, R.A., Simon, L.R.J., 1960. Two Studies in Automobile Insurance Ratemaking. ASTIN Bulletin 1(4), 192-217.
(9) Bayes, T., 1763. An Essay towards Solving a Problem in the Doctrine of Chances. Philosophical Transactions of the Royal Society of London 53, 310-418.
(10) Bühlmann, H., 1967. Experience Rating and Credibility. ASTIN Bulletin 4(3), 199-207.
(11) Bühlmann, H., 1970. Mathematical Methods in Risk Theory. Springer Verlang, New York.

(12) Bühlmann, H., Straub, E., 1970. Credibility for Loss Ratios. Bulletin of the Swiss Association of Actuaries 70, 111-133.
(13) Charpentier, A., Denuit, D., 2004. Mathématiques de l'Assurance Non-Vie, Tome I: Principe fondamentaux de théorie du risqué. Economica, Paris.
(14) Charpentier, A., Denuit, M., 2005. Mathématiques de l'Assurance Non-Vie, Tome II: Tarification et provisionnement. Economica, Paris.
(15) Cheng Hsiao, Changseob Kim, Grant taylor, A statistical perspective on insurance rate-making.
(16) Chiappori, P. A., Gollier, C. (Eds.). Asymmetric Information, in "Competitive Failures in Insurance Markets: Theory and Policy Implications." MIT Press, London, pp. 159-184.
(17) Chiappori, P. A., Jullien, B., Salanié, B., Salanié, F., 2006. Asymmetric Information in Insurance: General Testable Implications. RAND Journal of Economics 37, 783-798.
(18) Chiappori, P. A., Salanié, B., 2000. Testing for Asymmetric Information in Insurance Markets. Journal of Political Economy 108, 56-78.
(19) Cramér, H., 1930. On the Mathematical Theory of Risk. Skandia Jubilee, Swedish.
(20) Dean, C., J.Law lesset G.Willmot.1989. _ A mixed Poisson-inverse Gaussian regression model _. Canadian Journal of Statistics17,p.171-182.
(21) Delaporte, J.P., 1972. Les Mathématiques de l'Assurance Automobile. ASTIN Bulletin 6(3), 185-190.
(22) Denuit, M., 2003. Tarification Automobile sur Données de Panel, ASTIN Bulletin 36(1), 51-81.
(23) Denuit, M., 2006. An Actuarial Analysis of the French Bonus-Malus System. Scandinavian Actuarial Journal 2006(5), 247-264.
(24) Dionne, G., Gouriéroux, C., & Vanasse, C., 2001. Testing for Evidence of Adverse Selection in the Automobile Insurance Market: a Comment. Journal of Political Economy 109, 444-453.
(25) Filipe Charters de Azevedo, Teresa A. Oliveira, Amilcar Oliveira "Modeling Non-Life Insurance Price For Risk Without Historical Information" REVSTAT – Statistical Journal Volume 14, Number 2, April 2016, 171–192
(26) Frees, E. W., 2010. Regression Modeling with Actuarial and Financial Applications. Cambridge University Press, Cambridge.
(27) Frees, E. W., 2010. Regression Modeling with Actuarial and Financial Applications. Cambridge University Press, Cambridge.
(28) Gamadeku Mawuli, Application Of Bühlmanns Credibility Theory To An Automobile Insurance Claims.
(29) Gamadeku Mawuli, Application Of Bühlmanns Credibility Theory To An Automobile Insurance Claims.
(30) Gerder, H. U., 1979. An Introduction to Mathematical Risk Theory. S. S. Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania.
(31) Gómez-Déniz, E. (2009). Some Bayesian Credibility Premiums Obtained by Using Posterior Regret Gamma-Minimax Methodology. Bayesian Analysis, 4, 2, 223-242.
(32) Gouriéroux, C., Monfort, A.,& Trognon, A.(1984).Pseudo maximum likelihood methods: Theory.Econometrica 52,681-720.
(33) Jong, P., Zeller, G., 2008. Generalized Linear Models for Insurance Data. Cambridge University Press, Cambridge.
(34) Julien Antunes Mendes,Sébastien de Valeriola, Samuel Mahy ,Xavier Maréchal, (2017) Machine Learning applications Frequency modelling: An educational case study .

ISSN: 2658-8455
Volume 2, Issue 1 (January, 2021), pp. 344-361.
www.ijafame.org

(35) Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M., 2009. Modern Actuarial Risk Theory: Using R. 2nd ed. Springer, Berlin.
(36) Lemaire, J., 1995. Bonus-malus Systems in Automobile Insurance. Kluwer Academic Publishers, Boston.
(37) Lemaire, J.,1977,Selection procedures of regression analysis applied to automobile insurance, Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker,143-160.
(38) Lemaire, J.,2004, Bonus-Malus system, Encyclopedia of Actuarial Science, Wiley, New York,184-191.
(39) Lemaire, L., 1985. Automobile Insurance: Actuarial Models. Kluwer Academic, Dordrecht.
(40) Lotsi, A., Mettle, F. O., & Adjorlolo, P. K. (2019). Application of Bühlmanns-Straub Credibility Theory in Determining the Effect of Frequency-Severity on Credibility Premium Estimation. ADRRI Journal of Physical and Natural Sciences, 3(2), 1-24.
(41) Lundberg, E.F., 1903. Approximations of the probability function. Reinsurance of collective risks. Doctoral thesis, University of Uppsala, Swedish.
(42) M. SADOUN Ahmed ,méthode de tarification des risques en assurance non-vie, thèse.
(43) McClenahan, C.L., 2001. Ratemaking. 4th ed. Foundations of Casualty Actuarial Science, Casualty Actuarial Science.
(44) McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models. 2nd ed. Chapman and Hall, London.
(45) Mihaela David, A review of theoretical concepts and empirical literature of non-life insurance pricing, Procedia Economics, and Finance 20 (2015) 157 – 162.
(46) Mihaela David, Auto insurance premium calculation using generalized linear models Procedia Economics and Finance 20 (2015) 147 – 156.
(47) Mowbray, A.H., 1914. How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium. Proceedings of the Casualty Actuarial Society 1, 24-30.
(48) Nasser Hadidi, Classification Ratemaking Using Decision Trees.
(49) Ohlsson, E., Johansson, B., 2010. Non-Life Insurance Pricing with Generalized Linear Models. Springer, Berlin.
(50) Olivier Le Courtois. q-Credibility. Variance, 2020. Hal-02525182
(51) Payandeh Najafabadi, Amir T., 2010. "A new approach to the credibility formula," MPRA Paper 21587, University Library of Munich, Germany, revised 0020.
(52) Pesonen, E., 1962. A Numerical Method of Finding a Suitable Bonus Scale. ASTIN Bulleti, 2(1), 102-108.
(53) Ragnar Norberg, "Credibility Theory" Department of Statistics, London School of Economics, London WC2A 2AE, United Kingdom
(54) Savage, L.J., 1954. The Foundation of Statistics. Dover Publications, New York.
(55) Serpil Ergün Bülbül, Kemal B. Baykal, 2016. Optimal Bonus-Malus System Design in Motor Third-Party Liability Insurance in Turkey: Negative Binomial Model International Journal of Economics and Finance 8(8):205
(56) Spedicato, G., C. Dutang and L. Petrini, 2018, Machine Learning Methods to Perform Pricing Optimization. A Comparison with Standard GLMs. Variance 12, Issue 1:69-89.
(57) W. KIP VISCUSI, The Risky Business of Insurance Pricing. Journal of Risk and Uncertainty, 7:117-139 (1993)
(58) Whitney, A.W., 1918. The Theory of Experience Rating. Proceedings of the Casualty Actuarial Society 4, 274-292.
(59) Wüthrich, M., and C. Buser, 2019, Data Analytics for Non-Life Insurance Pricing.

(60) Zhang J., and T., Miljkovic, 2019, Ratemaking for a New Territory: Enhancing GLM Pricing Model with a Bayesian Analysis, Casualty Actuarial Society E-Forum, Spring 2018-Volume 2.

(61) Zhengmin Duan, Yonglian Chang, Wang, Tianyao Chen, and Qing Zhao. A Logistic Regression Based Auto Insurance Rate-Making Model Designed for the Insurance Rate Reform.