وزارة التعليم العالي والبحث العلمي

Université Badji Mokhtar
Annaba

Badji Mokhtar University -
Annaba

جامعة باجي مختار
عنابة

Faculté des Sciences

Département de Mathématiques

Laboratoire LaPS

THESE

Présentée en vue de l'obtention du diplôme de
Doctorat 3ème cycle en Mathématiques
Option :Statistique Actuarielle

Prime de crédibilité et méthode de l'entropie

Par:

METIRI FAROUK

Sous la direction de

ZEGHDOUDI Halim	M.C.A	U.B.M. Annaba

Co-directeur

REMITA Mohamed Riad	Prof.	U.B.M. Annaba

Devant le jury

PRESIDENT :	**BOUTABIA Hacène**	Prof.	U.B.M. Annaba
EXAMINATEUR :	**HADJI MohamedLakhdar**	M.C.A	U.B.M. Annaba
EXAMINATEUR :	**BRAHIMI Brahim**	M.C.A	U.M.K. Biskra

Année : 2016

Dedicace

To my parents

To my brother

To my sister

Acknowledgement

Firstly, I am extremely thankful to **Allah** who gave me chance and courage to complete this work.

I would like to express my sincere gratitude to my advisors Dr. ZeghdoudiHalimfor the continuous support of my Phd study and related research, for his patience, motivation. His guidance helped me in all the time of research and writing of this thesis.

Special appreciation to my co-supervisor Prof. Remita Mohamed Riad for his immense knowledge, his advices despite all the obligations that he had and his love of all mathematicians .

I would like also to thank the members of my doctoral committee: Prof. Boutabia Hacène, Dr. Hadji Mohamed Lakhdar and Dr. Brahimi Brahim, who agreed to serve on my committee on short notice and whose valuable comments will help improve this thesis.

I humbly and gratefully acknowledge the continual and enduring love and emotional support given by my parents and my brother and sister. I could never acknowledge or thank my parents enough.

I extend special thanks to Mr. Maroc Azzedine and Mr.Ferdjani Moncef for their encouragement to finish the thesis.

I would thank all my friends and colleagues and staffs of mathematics department and university, in particular “Ahmed”.

Special mention must be made here for my intimate friends who supported and motivated me to continue through the ups and downs towards my goal.

ملخص

في هذه الرسالة, نهتم بدراسة عنصر أساسي في نظرية المصداقية ألا و هو مقدر"**بيز**" لقسط التأمين.باعتبار توزيع" **ليندلاي**" كتوزيع خاص بمبالغ الأضرار و بالاعتماد على توزيعين أسبقيين للمقياس**θ** أحدهما معلوماتي و الأخر غير معلوماتي, نقوم باستخراج مقدر"**بيز**"تحت ثلاث دوال للخسارة:دالةالخطأ التربيعي التي هي دالة متناظرة و دالتان غير متناظرتين خاصتين بحساب الأخطاء و هما : **"linex, entropy".**

في هذه الحالة,مقدر"**بيز**" لقسط التأمين لا يملك شكلا خطيا مما يدفعنا لإستعمال تقنية عددية للتقريب و هي تقنية تقريب "**ليندلاي**".

قمنا بعمل محاكاة عددية لمقدر" **بيز**"بواسطة طريقة متوسط مربعات الخطأ **(MSE)**لتأكيد صحة النتائج النظرية المحصل عليها و مقارنة هذا المقدر تحت مختلف دوال الخسارة المذكورة أعلاه.

Contents

PERSONAL BIBLIOGRAPHY

Published articles

• Metiri, F. Zeghdoudi, H. Remita, M.R, On Bayes Estimates of Lindley Distribution under Linex Loss Function: Informative and Non Informative Priors. Global Journal of Pure and Applied Mathematics. ISSN 0973-1768, Volume 12, Number 1, pp. 4395-440 (2016).

• Metiri, F. Zeghdoudi, H. Remita, M.R, On weighted balanced loss function under the Esscher principle and credibility premiums. Hacettepe Journal of Mathematics and Statistics. ISSN 13035010. (Accepted).

Submitted article

• Metiri, F. Zeghdoudi, On Bayes Estimates of Lindley Distribution under Squared error and Linex Loss Functions: Informative and Non Informative Priors. Global Journal of Pure and Applied Mathematics. ISSN 0973-1768.

Résumé

Dans cette thèse, nous nous proposons, une synthèse sur la théorie de la crédibilité. Plus précisement, nous considérons la distribution de Lindley comme une distribution conditionnelle où nous dérivons l'estimateur de la prime bayésienne sous les fonctions de perte de l'erreur quadratique moyennne, linex et l'entropie avec des distributions a priori informatives et non informatives en utilisant l'approximation de Lindley. Une simulation numérique et une étude comparative sont obtenues.

Mots clés: Distribution de Lindley, Fonction de perte de l'entropie, Méthode des moindres carrées, Prime Bayesienne, Théorie de la Crédibilité, .

Abstract

In this thesis, we focuse on a popular tool in credibility theory which is the Bayesian premium estimator, considering Lindley distribution as a claim distribution, we derive this estimator under entropy, squared error and linex loss functions with informative and non-informative priors.

The Bayesian premium estimator which is non linear in this case is solved using numerical approximation (Lindley approximation). Simulation study is then performed to evaluate this estimator and Mean squared error is computed to compare it under different loss functions.

Key words: Bayesian premium, Credibility theory, Entropy loss function, Lindley distribution, Mean squared error.

Introduction

L'assurance est une opération par laquelle une personne, l'assureur, s'engage à exécuter une prestation au profit d'une autre personne -l'assuré- en cas de réalisation d'un événement aléatoire, le risque (ou sinistre), en contrepartie du paiement d'une somme, cette somme est appelée la **prime** ou la **cotisation**.

L'actuaire travaille au sein des compagnies d'assurance. Sa mission: estimer les réserves que la compagnie doit établir pour faire face aux dépenses en fonction des différents contrats signés avec ses clients, calculer les risques et prévoir tous les aléas d'une situation donnée, de manière à faire une tarification qui ne laisse aucune place à l'imprévu et qui maximise les bénéfices de la compagnie.

Le calcul de la prime pure est un point fondamental en tarification, il a pour but d'évaluer pour chaque assuré le montant attendu des sinistres pour la période d'assurance étudiée. Ce calcul se fait le plus fréquemment par des méthodes statistiques comme: les modèles linéaires généralisées (GLM), bonus-malus et experience rating.

La tarification basée sur l'expérience (experience rating) en assurance est une technique qui vise à assigner à chaque risque sa prime juste et équitable. Cette prime pour une période dépend exclusivement de la distribution des sinistres de ce risque pour cette période. Cette technique exige un volume d'expérience important. Elle est donc principalement utilisée en assurance automobile et en accidents du travail.

Elle ne peut toutefois être utilisée, par exemple, en assurance-vie (on ne meurt qu'une fois) ou en assurance habitation (fréquence trop faible).

Développée par les écoles suisse et scandinave, la théorie de la crédibilité qui est le pilier d'expérience rating repose sur les principes de l'inférence Bayesienne et représente un ensemble des techniques utilisées par les actuaires pour déterminer la prime d'un assuré/contrat dans un portefeuille hétérogène en utilisant les informations des années précédentes. Elle s'est développée en parallèle de la statistique Bayesienne.

Pour un assuré, son risque X est caractérisé par une réalisation notée θ. Chaque sinistre est vu donc comme une variable aléatoire selon une distribution conditionnelle qui dépend de θ. En plus, l'espérance des sinistres appelée aussi la **prime de risque (individuelle)** peut être calculée sur la base de cette distribution conditionnelle.

En pratique, différents modèles ont été proposés pour chercher le meilleur estimateur de la prime de risque. Le modèle Bayesien quant à lui, consiste à supposer pour θ une distribution des sinistres à priori $\pi(\theta)$ -appelée aussi fonction de structure- qui décrit l'expérience de risque du même paramètre appelée $f(x \mid \theta)$.

En combinant la distribution à priori avec la vraisemblancee de $f(x \mid \theta)$,on peut obtenir la distribution à posteriori $f(\theta \mid x)$ qui établit la dépendance de θ sachant l'historique de l'expérience. La prime Bayesienne P^B -meilleure prime d'expérience- représentant l'espérance des sinistres futurs peut être calculée à partir de $f(\theta \mid x)$.

Cependant, la prime Bayesienne qui sera chargée à l'assuré est une prime de crédibilité - ayant une forme linéaire- seulement sous une famille de distributions et lois à priori conjuguées spécifiées et aussi sous la fonction de perte de l'erreur quadratique.

$$\mathbf{P^{B}} = \mathbf{z} \times \mathbf{P^{experience}} + (\mathbf{1} - \mathbf{z}) \times \mathbf{P^{collective}} \tag{1}$$

On s'intéresse à une autre fonction de perte qui attribue plus de poids à la surestimation. Dans la théorie de la décision, la fonction de perte de l'entropie $\left(L(\theta,\hat{\theta}) = \left(\frac{\hat{\theta}}{\theta}\right)^{q} - qln\left(\frac{\hat{\theta}}{\theta}\right) - 1,\ q > 0\right)$ est une fonction qui donne une importance à la surestimation plus considérablement que la sousestimation.

La forme de cet estimateur dans le cas ou $\pi(\theta)$ n'est pas conjuguée est difficile à obtenir à cause des intégrales compliquées.

Pour pallier à ce problème et dériver la prime Bayesienne, on utilise l'approximation de Lindley qui est une méthode numérique d'approximation très utilisée pour résoudre ces formes d'intégrales et qui donne des résultats numériques.

La thèse s'articule autour de trois chapitres, le premier chapitre est consacré aux rappels d'historique et à la formulation mathématique nécessaire de la théorie de la crédibilité. En chapitre 2, on aborde les propriétés de la fonction de perte de l'entropie et on cherche l'expression de l'estimateur de la prime Bayesienne sous cette fonction de perte.

Le troisième chapitre est dédié à l'estimation de la prime Bayesienne dans le cas ou il est impossible d'établir une prime de crédibilité, on utilise une combinaison entre la distribution de Lindley et des lois à priori informatives et non informatives (l'extension de Jeffrey et inverse gamma) sous trois fonctions de perte (erreur quadratique , linex et entropie). On termine ce chapitre par une simulation numérique et une étude comparative de cet estimateur obtenu sous différentes fonctions de perte. Les conclusions et les perspectives sont données dans le quatrième chapitre.

Introduction

Experience rating is one of the most important practices in pricing for insurers. Insurance companies set up experience rating systems to determine pricing of premiums for different groups or individuals based on their past experience. In a competitive market nowadays, insurers want to determine individual premiums as precisely as possible.

Individual policyholders are usually divided into different groups according to their deemed "risk levels", which are often assessed during the underwriting process based on a variety of relevant factors. A manual rate is then introduced for each group to represent the expected experience arising from the unique risk characteristics of the class.

One implicit assumption embedded in the manual rate is that the underlying risk level is uniformly the same for each member of the class, which is sometimes referred to as "homogeneity" by actuaries. However, as [9] pointed out, there are actually no homogeneous risk classes in insurance. Empirical evidence has suggested that individual experiences may vary considerably even within the same risk group, because no risk is exactly the same as another. [46] also point out that such heterogeneity may only appear to the insurer through the individual claims records. Therefore, insurance premiums cannot be solely determined by manual rates. Unique individual experiences also need to be taken into account.

Credibility theory is an experience rating technique in actuarial science which can be seen as one of quantitative tools that allows the insurers to perform experience rating, that is, to adjust future premiums based on past experiences. It is used usually in automobile insurance, worker's compensation premium, loss reserving and IBNR (Incurred But Not Reported claims to the insurer) where credibility theory can be used to estimate the claim size amount.

In this sense, credibility theory is used to determine the expected claims experience of an individual risk when those risks are not homogeneous, given that the individual risk belongs to a heterogeneous collective. The main objective of this theory is to calculate the weight which should be assigned to the individual risk data to determine a fair premium to be charged. Introduction to credibility theory can be found, e.g., in [25], [29], [33], [34], and for recent detailed introductions, one can refer to, for example, [46], which describes its evolutionary history and gives a simple account of its main issues and results, and [9], which gives a comprehensive exposition of the modern credibility theory.

In credibility theory, each risk X of an insured is characterized by a distribution identified by an unknown risk parameter, and due to the heterogeneity over policies in the concerned portfolio, all possible values of θ are modeled with certain random variable following a probability distribution $\pi(\theta)$, which is referred to a structure function in actuarial context and prior distribution in statistical theory. To predict a possible future loss of the risk X or estimate its mean by a quantity (referred to as risk or Bayesian premium P^B), one observes a sequence of its historical claims and then effectively summarizes information from the observed data (referred to as experience below). Moreover, to the assumed distribution $\pi(\theta)$ one can associate a premium (referred to as collective premium below). Consequently, P^B is represented by a credibility form

$$\mathbf{P^B} = \mathbf{z} \times \mathbf{P^{experience}} + (\mathbf{1} - \mathbf{z}) \times \mathbf{P^{collective}} \tag{2}$$

However, the credibility premium presented above is restricted by a family of distributions and conjugate prior. Neither the claim distributions which are not members of the exponential family of distributions nor the non-conjugate prior, the predicted mean (Bayesian premium estimator with respect to square error loss) is no longer linear with respect to the data (see [16]) and the credibility formula is no longer true. Whenever the policyholder is undercharged (and insurance company loses its money) or the insured is overcharged (and the insurer is at risk of losing the policy), the square loss assigns similar penalty to over and undercharge. In order to assign more (or less) penalty to overcharged, one has to consider a nonnegative convex function as a loss function rather than square error loss to reflect such concerns.

In this case, we may be interested in a loss function which assigned more penalty to overcharges. In decision theory, entropy loss function (given by $L(\theta, \hat{\theta}) = \left(\frac{\hat{\theta}}{\theta}\right)^q - qln\left(\frac{\hat{\theta}}{\theta}\right) - 1$, $q > 0$) is a popular loss which consider in situation that overestimation is more considerable than underestimation.

It is well known that the Bayesian premium enjoys the advantage of being the estimator with the least squared error loss. However, the explicit form of this estimator in the case when considering $\pi(\theta)$ not member of the exponential family could be quite difficult to obtain as it involves a number of integrations which are not analytically solvable.

Therefore, one has to use an approximation method for the solutions. Lindley's approximation technique is one of the methods suitable for solving such problems which approaches the ratio of the integrals as a whole and produces a single numerical result.

In this thesis, our focus will be restricted to determination of Bayesian premium within greatest accuracy credibility theory. One important aspect for this model is the choice of the distribution of losses, which depicts the pattern of the experience of a policyholder conditional upon his risk parameter value. Empirical evidence often suggests that the individual's claim data can sometimes be volatile and hard to predict.

This thesis is organized as follows. The mathematical formulations underlying the history of credibility theory including Bayesian premium are carefully reviewed in Chapter 1. In Chapter 2, we describe the entropy loss and its properties, we present the expression of the Bayesian premium under this loss when this latter is considered as credibility formula, i.e. under the exponential family. The explicit solution of the Bayesian premium estimator is then obtained.

The Chapter 3 presents our result, we establish the estimation of the Bayesian premium in the case where there is no credibility formula by using a combination between lindley distribution and informative and non informative priors.

The Bayesian premium estimator is treated under three loss functions (squared error, linex and entropy) using informative and non informative priors (the extension of jeffrey and the inverted gamma, respectively). Simulation study will be performed and comparisons will be conducted accordingly. Conclusions and future work are presented in Chapter 4.

Chapter 1

Historical Review Of Credibility Theory

Credibility theory uses two main approaches, each representing a different method of incorporating individual experience in the ratemaking process. The first and oldest approach is called **limited fluctuation credibility** (also referred to *American credibility*). According to this approach, an insured's premium should be based solely on its own experience if the experience is significant and stable enough to be considered credible. The second is called **greatest accuracy credibility** (also referred to as *European credibility*).

In the first chapter, we present the credibility theory which is a set of quantitative tools used by insurers for performing experience ratings. Two branches of credibility theory, known as limited fluctuation credibility theory and greatest accuracy credibility theory have also been discussed, we will provide more detailed discussions around mathematical assumptions and formulations of these two approaches.

1.1 Limited fluctuation credibility theory

The birth of credibility theory dates back to the beginning of the century with a paper by [43] "How extensive a payroll exposure is necessary to give a dependable pure premium?". In the workers compensation insurance field, *Mowbray* was interested in finding the minimal number of employees covered by a plan such that the premium of the employer could be considered fully dependable, that is, fully *credible*.

An individual insured's premium to be fully credible if it fluctuates moderately from one period to another. That is, the credibility criterion is stability. of experience, which usually increases with the volume of the insured's experience. This volume is expressed as number of claims, number of employees, square foot of factory surface, etc.

With the emergence of theoretic methods, *Mowbray*'s original problem can be formulated in a slightly more general way as follows.

Let us define the random variables:

N_t = Number of claims generated by the insured during the t^{th} time period (months, quarters, years, etc.), for $t = 1, 2, ...,$

X_{tj} = Size of the j^{th} claim in the t^{th} year, for $j = 1, 2, ..., N_t,$

S_t = Size of the aggregate claims in the t^{th} period of time.

Then,

$$S_t \quad = \sum_{j=1}^{N_t} X_{tj} = X_{t1} + X_{t2} + ... + X_{tN_t} \ ,$$

where X_{tj}s are assumed to be independent, identically distributed (i.i.d) random variables that are also mutually independent of the N_ts. This is the collective model of risk theory. Most of the situations usually encountered in limited fluctuation credibility can be described by an application of this model. It is also well-known (see, for example, [21]) that

$$\begin{cases} E\left[S_t\right] = \ E\left[N_t\right] E\left[X_{tj}\right] \\ Var\left[S_t\right] = \ E\left[N_t\right] Var\left[X_{tj}\right] + Var\left[N_t\right] E\left[X_{tj}\right]^2 \end{cases} .$$

Let $\bar{S}_t = \frac{(S_1+S_2+......S_T)}{T}$ denote the insured's observed average (empirical mean) claim amount at the end of T periods, $T = 1, 2, ...$

The fundamental problem of limited fluctuation credibility is the determination of the parameters of the distribution of $\bar{S}_t$ such it stays within $100k$ percent of its expected value with probability p, i.e.,

$$Pr[(1-k)E[\bar{S}_t] \leq \bar{S}_t \leq (1+k)E[\bar{S}_t]] \geq \ p, \tag{1.1}$$

holds for given p and k. In a typical limited fluctuation credibility situation, the parameter k is small $(e.g., \ 5 \text{ to } 10 \text{ percent})$, while parameter p is large $(\text{often above } 90 \text{ percent})$.

1.1.1 Full credibility

When an insured meets the requirements of (1.1), the insured is said to deserve a full credibility of order $(k,\ p)$, i.e., the insured is charged a pure premium based solely on the insured's own claims experience. If full credibility occurs after T^*periods, the credibility premium would be $\bar{S}_{T^*}$.

Equation (1.1) thus requires the distribution of $\bar{S}_T$ to be relatively concentrated around its mean. As $\bar{S}_T$ is a sum of i.i.d random variables, the distribution of $\bar{S}_T$ has to be approximated.

Assuming the second moment of $\bar{S}_T$ is finite, one can use the version of the central limit theorem applicable to random sums (see [18], p.258) to approximate the distribution. Thus:

$$\frac{\bar{S}_T - E[\bar{S}_T]}{\sqrt{VAR\left[\bar{S}_T\right]}} \rightarrow^{n\rightarrow\infty} N(0,1),$$

i.e., a standard normal distribution. (1.1) may then be rewritten

$$Pr[\frac{-kE[\bar{S}_t]}{\sqrt{VAR\left[\bar{S}_t\right]}} \leq \frac{S - E[\bar{S}_t]}{\sqrt{VAR[\bar{S}_t]}} \leq \frac{kE[\bar{S}_{\mathfrak{t}}]}{\sqrt{VAR\left[\bar{S}_t\right]}}] \approx 2\Phi[\frac{kE[\bar{S}_t]}{\sqrt{VAR\left[\bar{S}_t\right]}}] - 1 \geq p,$$

Hence

$$\left(E[\bar{S}_t]\right)^2 \geq \left(\frac{\zeta_{1-\varepsilon/2}}{k}\right)^2 \frac{Var\left[S_t\right]}{T}, \tag{1.2}$$

Where $\varepsilon = 1 - \rho$ and $\boldsymbol{\zeta}_\alpha$ is the α^{th} percentile of a standard normal distribution.

At this point, the essence of the theory of limited fluctuation credibility (i.e., equation (1.1)) has been covered.

The *Mowbray*'s solution provided just that, a level above which an individual premium is granted full credibility and zero credibility below that level. However, an insured with total number of claims just below the full credibility level may pay a significantly different premium.

The dichotomy between zero and full credibility paved the way for the development of partial credibility.

1.1.2 Partial credibility

The first partial credibility formula was developed by *Albert Withney* In his 1918 paper, *Withney* refers to "necessity, from the standpoint of equity to the individual risk, of striking a balance between class-experience on the one hand and risk-experience on the other". The objective of credibility theory is the calculation of this balance.

Which principles should govern the calculation of this balance?, According to *Withney* , the balance depends on four elements: the exposure, the hazard, the collective premium, and the degree of concentration within the class (homogeneity of the entire portfolio).

After some calculations, he obtains the following expression for the individual's premium P :

$$P = z\bar{X} + (1 - z)\, m, \tag{1.3}$$

Where $\bar{X}$ is the mean from the individual's experience and m is the collective mean. Notice that $\bar{X}$ and m are combined to produce a weighted average with z and $z-1$ as weights. An expression of the form of (1.3) is called a *credibility premium* . The quantity z is called *the credibility factor* which is given by

$$z = \frac{n}{n+k},$$

Note that k is not an arbitrary constant, rather it is an explicit expression that depends on the various parameters of the model. For the sake of simplicity and to avoid large fluctuations between the individual and collective premiums, however, Withney suggests that k may be determined by the actuary's judgement rather than by its correct mathematical formula.

1.1.3 Uses of limited fluctuation credibility

From a theoretical perspective, the range of applications of limited fluctuation credibility is fairly limited, though many of these are ignored in practice. the key point to remember when using limited fluctuation credibility is that it relies solely on stability criterion, which, generally, is the size of insureds or the number of periods (years quarters, etc) of claims experience. As such, limited fluctuation credibility should be used only when stability of the experience is of foremost importance.

The case for partial credibility has been successfully used by American actuaries to restrict premium variation from one time period to another. One can argue that partial credibility takes into account the heterogeneity of the insurer's block insureds by charging different premiums to different groups of insureds. This differentiation among the insureds, however, is only based on their size or the extent of their claims history, this is not necessarily fair.

One must bear in mind that the goal of partial limited fluctuation credibility is not to calculate the most precise premium for an insured. The goal is to incorporate into the premium as much individual experience as possible while still keeping the premium sufficiently stable. It is important to understand this distinction. When credibility is used to find the most precise estimate of an insured's pure risk premium, one must turn to **greatest accuracy credibility** methods.

1.2 Greatest accuracy credibility

Greatest accuracy credibility is a more modern, versatile, and complex field of credibility theory. It is not a single theory, rather it is an approach to the credibility problem. The main objective is to find the best premium to charge an insured, where best is in the sense that the premium estimator is the closest estimator to the true premium.

One important point to keep in mind when moving from limited fluctuation to greatest accuracy credibility is that a high credibility factor (i.e., z close to 1) is no longer a goal in itself. Indeed, the credibility factor will henceforth mostly reflect the degree of heterogeneity of the portfolio, rather than the degree of stability of an individual's risk experience. For an homogeneous portfolio, greatest accuracy credibility states there is to need to charge a different premium to the insureds. The credibility factor will accordingly be low, i.e., close to 0. Conversely, the more heterogeneous portfolio, the greater the consideration of the individual experience, hence the higher the credibility factor.

To illustrate this, imagine a portfolio consisting of five very large insureds, each having identical means. Given the importance of their size, each group of insureds would all be granted full credibility under the limited fluctuation approach. As their means are all equal, however, they form a perfectly homogeneous portfolio. Accordingly, their credibility level will be zero under the greatest accuracy approach. Of course, the end result is the same because the collective mean is equal to the individual means, but this shows how different can be the interpretation of the credibility factor in greatest accuracy credibility.

1.2.1 The mathematical model

Consider an insurance portfolio consisting of I insureds. The ideal situation for ratemaking occurs if this portfolio is relatively homogeneous, i.e., the insureds have similar risks characteristics of insured i that reflects the insured's risk level which is denoted by the risk parameter θ_i for insured $i = 1, 2, ..., I$. This risk parameter incorporates every characteristic of the insured that is not otherwise accounted for in the initial risk classification process.

The parameter θ_i is unknown and is assumed to be constant throughout the life of the insurance contract. Because of the assumption of an homogeneous portfolio, we must further assume that each insured's θ_i is viewed as being drawn at random from the same cumulative distribution, following [8], $\pi(\theta)$ is called the *structure* or *prior function*.

In a purely Bayesian setting, $\pi(\theta)$ represents the insurer's *prior* belief about the insured's risk level. After collection of the insured's data at the end of the period, the insurer's initial judgement is revised and the structure function modified accordingly. This interpretation is particularly suited to the case where there is a single insured or when the insurer has little information and must take an educated guess at the initial pure premium for example, when the insurer is entering a new line of business where no data are available.

In the rest of this thesis, we consider the purely Bayesian setting with only one insured (so the subscript i will be dropped). The claim amounts X_t $(t = 1, 2, ...)$ are independent and identically distributed, but only given θ, the risk parameter of the insured. Unconditionally, the X_ts are not necessarily independent. The conditional distribution of $X \mid \theta$ is denoted by $f(x \mid \theta)$.

The determination of a claim amount can be demonstrated with a two-urn model. The first urn represents the urn containing the collective with distribution function $\pi(\theta)$. From this urn, we select the individual risk, or equivalently its risk level θ. The parameter θ then determines the content of the second urn, or equivalently the conditional distribution $f(x \mid \theta)$. From this urn, we select the values of the random variables $X_1,\ X_2, ...,$

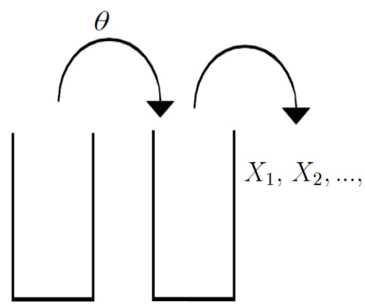

Figure.1
$\pi(\theta)$ and $f(x \mid \theta)$

1.2.2 Definition of various premiums

An underlying tenet of credibility theory is that the estimated premium is the pure or net premium, without any provision for random fluctuations, profits, or expenses. Thus two insureds with different variances but the same mean are charged the same pure premium.

We distinguish here between four types of **pure** premiums: the individual premium, the collective premium, the Bayesian premium, and the credibility premium.

Individual premium (risk premium)

Definition 1.1 *The risk premium, $\mu(\theta)$, is the correct individual premium to charge an insured if the insured's risk level, θ, is known.*

The risk premium is thus the expected value of the insured's claim amount in one period, given his or her risk level.

It is given by

$$\mu(\theta) = E[X \mid \theta] = \int_0^\infty x f(x \mid \theta)\, dx. \tag{1.4}$$

The individual premium is also referred to as the fair risk premium. Because the risk parameter θ is unobservable in practice, $\mu(\theta)$ can never be exactly known and hence must be estimated from data.

Collective premium

An insurance company insures many kinds of risks. For the purpose of rating risks, these risks are grouped into classes of "similar risks". In motor insurance, examples of such risk characteristics are cylinder capacity, make of car and power/weight ratio, as well as individual characteristics such as the driver's age, sex and region. In industrial fire insurance, important characteristics might be the type of construction of the insured building, the kind of business conducted in the building, or the fire extinguishing facilities in the building.

The most important thing for the credibility theory framework is that we do not consider each risk individually but that we rather consider each risk as being embedded in a group of "similar" risks, called the collective.

Definition 1.2 *The collective premium, m, is the pure premium charged when nothing is known about the insured's risk level (during the first year, for example). The collective premium is in essence the average value of all possible risk premiums. Math-*

ematically, the collective premium is given by

$$m = E[X\ ] = E[E[X\ \mid \theta]\ ] = E[\mu(\theta)]. \tag{1.5}$$

The fundamental difference between limited fluctuation and greatest accuracy credibility is the type of estimator of the risk premium. In limited fluctuation credibility, the observed claim average $\bar{X}$ is chosen if the experience is sufficiently stable and fully credible, otherwise, the collective mean m is charged. On the other hand, the objective in greatest accuracy credibility is to find an estimator as close as possible to the true value of $\mu(\theta)$ given the available data. There is no unique way to measure closeness. In Bayesian credibility, for example, the most used closeness measure is the mean square error between the estimator and the individual premium.

Bayesian premium (best experience premium)

Definition 1.3 *Suppose the data for* n *consecutive periods are* $X_1, ..., X_n = \underset{\sim}{X}$ *, then the Bayesian premium* P^B *is given by*

$$\mathrm{P}^B = \min_{g(.)} E\left[(\mu(\theta) - g(\underset{\sim}{X}))^2\right], \tag{1.6}$$

where $g(.)$ *is some function of the data.*

Under the squared error loss, It is no difficult to prove that the solution to this minimization problem is

$$\begin{aligned} \mathrm{P}^{B} &= E\left[\mu(\theta) \mid \underline{\mathrm{X}}\right]. \\ &= \int_{0}^{\infty} \mu(\theta) f(\theta \mid \underline{\mathrm{X}})\, d\theta. \end{aligned}$$

The Bayesian premium can thus be calculated in two steps:

1. First, by calculating the posterior distribution of θ given the data, $f(\theta \mid x_1, ..., \ x_n)$. Recall that the conditional distribution of $X \mid \theta$ and the distribution of θ are assumed to be known in the present model. From Bayes theorem and the conditional independence of claim amounts, we have

$$\begin{aligned} f(\theta / x_1, ..., \ x_n) &= \frac{f(x_1, ..., x_n \ \mid \ \theta)\, \pi(\theta)}{\int_\theta f(x_1, ..., x_n \ \mid \ \theta)\, d\theta} \\ &= \frac{\Pi_{j=1}^{n} \ f(x_j \ \mid \theta)\, \pi(\theta)}{\int_\theta \Pi_{j=1}^{n} \ f(x_j \ \mid \theta)\, d\theta} \propto \pi(\theta)\, \Pi_{j=1}^{n} f(x_j \ \mid \ \theta). \end{aligned} \tag{1.7}$$

2. Then, by calculating the expected value of $\mu(\theta)$ with respect to this distribution:

$$E\left[\mu(\theta) \mid \underline{\mathrm{X}}\right] = \int_\theta \mu(\theta) f(\theta \mid \underline{\mathrm{X}})\, d\theta. \tag{1.8}$$

where $\propto$ is the proportionality operator, i.e., the right-hand side is equal to the left-hand side up to a multiplicative constant not depending on θ. Calculation of the expected value is then immediate.

Here, it should be noted that the experience premium is a random variable (it is a function of the observation vector $\underset{\sim}{X}$), the values of which are known at the time at which the risk is to be rated. The Bayesian premium is thus a "true" premium, the value of which depends on the claim experience.

The last premium to be defined before we turn to exact Bayesian credibility is the credibility premium.

Credibility premium

We have seen that the Bayes premium $\mathrm{P}^{B} = E\left[\mu\left(\theta\right) \mid \underset{\sim}{X}\right]$ is the best possible estimator in the class of all estimator functions. In general, however, this estimator cannot be expressed in a closed analytical form and can only be calculated by numerical procedures. Therefore it does not fulfil the requirement of simplicity. Moreover, to calculate P^{B}, one has to specify the conditional distribution as well as the a prior distribution, which, in practice, can often neither be inferred from data nor guessed by intuition.

The basic idea underlying credibility is to force the required simplicity of the estimator by *restricting* the class of allowable estimator functions to those which are *linear* in the observations $X_1, ..., X_n$. In other words, we look for the best estimator in the class of all *linear estimator functions*.

"Best" is to be understood in the Bayesian sense and the optimally criterion is again quadratic loss. Credibility estimators are therefore linear Bayesian estimators.

A credibility premium P *, is a linear function of a special type of observations* $X_1, ..., X_n$ *of an insured, it is a convex combination between the individual experience weighted average* $\bar{X}$ *and the collective premium* m*, i.e.,*

$$P(X_1, ..., X_n) = z\bar{X} + (1 - z)m, \tag{1.9}$$

where $0 \leq z \leq 1$ *is the credibility factor and* $(1 - z)$ *is the complement of credibility.*

It should be noted that the complement of credibility is given to the collective premium, m, and nothing else.

1.2.3 Exact Bayesian credibility

To an actuary who considers himself to be a Bayesian, the Bayesian premium equation $\left(\mathrm{P}^B = E\left[\mu(\theta) \mid \underline{\mathrm{X}}\right]\right)$ is the best premium (in the least square sense) to charge an insured considering the experience at hand.

The Bayesian premium, however, has some drawbacks when it comes to being used in practice, the actual distributions of $X \mid \theta$ and θ must be known.

Moreover, unlike a credibility premium, there is no guarantee that a Bayesian premium will lie between the individual experience average and the collective premium. This fact can be difficult to explain to a layperson.

Fortunately, there are some combinations of distributions where the Bayesian premium has a nice form. Actually, in these cases, Bayesian premiums are exact credibility premiums.

We present in below an example of a combination in which, P^B is a compromise between $\bar{X}$ and m.

Example 1.1

suppose $X \mid \theta$ has a Poisson distribution with parameter θ, and θ has a gamma distribution of parameters α and β, i.e.,

$$f(x \mid \theta) = \frac{\theta^x e^{-\theta}}{x!}, x = 0, 1, ...$$

and

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \quad \theta > 0, \alpha > 0, \beta > 0.$$

The individual premium is

$$\mu(\theta) = E[X \mid \theta] = \theta.$$

Consequently, the collective premium is

$$m = E[\theta] = \frac{\alpha}{\beta}.$$

We find that the posterior distribution of θ is:

$$\begin{aligned} f(\theta \mid x_1, ..., \ x_n) &\propto \theta^{\alpha-1} e^{-\beta\theta} \Pi_{i=1}^n \theta^{x_i} e^{-\theta}, \\ &\propto \theta^{\alpha+\sum_i x_i - 1} e^{-(\beta+n)\theta} \ , \end{aligned}$$

which is a gamma distribution with updated parameters $\tilde{\alpha} = \alpha + \sum_i x_i$ and $\tilde{\beta} = \beta + n$, where $\underset{\sim}{X} = X_1, ..., X_n$. The Bayesian premium is thus

$$\begin{aligned}
\mathrm{P}^B &= E\left(\mu\left(\theta\right) \mid \underset{\sim}{X}\right), \\
&= E\left(\theta \mid \underset{\sim}{X}\right) = \frac{\tilde{\alpha}}{\tilde{\beta}}, \\
&= \frac{\alpha + \sum_i x_i}{n + \beta}, \\
&= \frac{\alpha + n\bar{x}}{n + \beta}, \\
&= \frac{n}{n+\beta}\bar{x} + \left(1 - \frac{n}{n+\beta}\right)\frac{\alpha}{\beta}, \\
&= z\bar{X} + (1 - z)\, m,
\end{aligned}$$

with $z = \frac{n}{n+\beta}$. The Bayesian premium is a convex combination between the individual experience average and the collective premium, i.e., a credibility premium with credibility factor z.

[3] was one of the first to show that for some combinations of distributions the Bayesian estimator is exactly a (linear) credibility premium. In doing so, *Bailey* also provided the exact value of the constant k in the credibility factor that [55] choose to determine by judgement. A few years later, [42] extended *Bailey*'s results.

The other combinations of distributions known to yield exact credibility premiums are presented in table (1.1).

$f(x \mid \theta)$	Bernoulli(θ)	Geometric(θ)	Exponential(θ)	$\mathrm{N}(\theta, \sigma_1^2)$
$\pi(\theta)$	Beta(α, β)	Beta(α, β)	Gamma(α, β)	$\mathrm{N}(\mu, \sigma_2^2)$
$f(\theta \mid \underline{X})$	$\mathrm{Beta}\left(\tilde{\alpha}_1, \tilde{\beta}_1\right)$	$\mathrm{Beta}\left(\hat{\alpha}_2, \hat{\beta}_2\right)$	$\mathrm{Gamma}\left(\tilde{\alpha}_3, \tilde{\beta}_3\right)$	$\mathrm{N}\left(\tilde{\mu}, \tilde{\sigma}_2^2\right)$
$\mu(\theta)$	θ	$\frac{1-\theta}{\theta}$	$\frac{1}{\theta}$	θ
m	$\frac{\alpha}{\alpha+\beta}$	$\frac{\beta}{\alpha-1}$	$\frac{\beta}{\alpha-1}$	μ
P^B	$\frac{\alpha+\sum_i x_i}{\alpha+\beta+n}$	$\frac{\beta+\sum_i x_i}{\alpha+n-1}$	$\frac{\beta+\sum_i x_i}{\alpha+n-1}$	$\frac{n\sigma_2^2\bar{x} + \sigma_1^2\mu}{n\sigma_2^2 + \sigma_1^2}$
z	$\frac{n}{n+\alpha+\beta}$	$\frac{n}{n+\alpha-1}$	$\frac{n}{n+\alpha-1}$	$\frac{n}{n+\frac{\sigma_1^2}{\sigma_2^2}}$

Table 1.1 - Bayesian credibility models for certain conjugate distributions pairs

$\left(\tilde{\alpha}_1, \tilde{\beta}_1\right)$	$\left(\alpha + \sum_j x_j, \beta + n - \sum_i x_j\right)$
$\left(\tilde{\alpha}_2, \tilde{\beta}_2\right)$	$(\alpha + n, \beta + \sum_i x_j)$
$\left(\tilde{\alpha}_3, \tilde{\beta}_3\right)$	$(\alpha + n, \beta + \sum_i x_j)$
$(\tilde{\mu},\ \tilde{\sigma}_2^2)$	$\left(\frac{\sigma_2^2 \sum_j x_j + \sigma_1^2\mu}{n\sigma_2^2 + \sigma_1^2}, \frac{\sigma_1^2\sigma_2^2}{T\sigma_2^2 + \sigma_1^2}\right)$

Table 1.2 - New parameters of the posterior distribution

The table (1.1) contains the distributions which are members of the so-called *exponential family*. [32] unified the results of this table in an elegant way.

Goel in [24] conjectured that only combinations of unidimensional exponential family members with their *natural conjugate* priors yield linear Bayesian premiums. If *Goel* is correct, then the only Bayesian premiums that are exact credibility premiums are the ones found in the Table 1.1 above.

1.3 The Bühlmann Model

We have two practical problems is Bayesian credibility:

- The Bayesian premium is a credibility premium in certain cases only,
- The premium on subjective assumptions for the distributions of θ_i and $X_i \mid \theta_i$.

So far we have only considered one particular risk and we have derived the credibility estimator based only on the observations of this particular risk. In practice, however, one usually has observations of a whole portfolio of similar risks numbered $i = 1, 2, ..., I$.

We denote by $X = (X_{11}, ..., X_{In})$ the observation vector of risk with $i = 1, 2, ..., I$ and $t = 1, 2, ..., n$. The portfolio is composed of I insureds each characterized by an observable random risk parameter θ_i (risk profile). We now assume that for each risk in the portfolio, θ_i and X_i fulfil the assumptions of the simple credibility model, i.e. θ_i and X_i are the outcomes of the two-urn model described above, applied independently to all risks i. We then arrive at the simple Bühlmann model published in the seminal paper "Experience Rating and Credibility" (see [7]).

1.3.1 Model Assumptions

$(\mathbf{B}_1)$: The random variables are, conditional on θ_i, independent with the same distribution function $f\ (x \mid \theta)$ and the conditional moments

$$\mu(\theta) = E[X_{ij} \mid \theta_i],$$

$$\sigma^2(\theta) = Var[X_{ij} \mid \theta_i],$$

$(\mathbf{B}_2)$: The pairs $(\theta_1, X_1), ..., (\theta_I, X_n)$ are independent and identically distributed.

Remark 1.2

- By Assumption $(\mathbf{B}_2)$, an heterogeneous portfolio is modelled. The risk profiles $\theta_1, \theta_2, ..., \theta_I$ are independent random variables drawn from the same urn with structural distribution $\pi(\theta)$. Hence, the risks in the portfolio have different risk profiles (heterogeneity of the portfolio). On the other hand, the risks in the portfolio have something in common: a priori they are equal, i.e. a priori they cannot be recognized as being different.

We now want to find the credibility estimator in the simple Bühlmann model. i.e., we want to estimate for each risk i its individual premium $\mu(\theta_i)$. Hence, there is not just one credibility estimator, but rather we want to find the credibility estimators of $\mu(\theta_i)$ for $i = 1, 2, ..., I$. By definition, these credibility estimators $\Pi(\theta_i)$ have to be linear functions of the observations. But here the second question arises: linear in what? Should the credibility estimator of $\mu(\theta_i)$ just be a linear function of the observations of risk i, or should we allow for linear functions of all observations in the portfolio?. Hence, there are always two items to be specified with credibility estimators: the quantity that we want to estimate and the statistics that the credibility estimator should be based on.

Generally, the credibility estimator is defined as the best estimator which is a linear function of all observations in the portfolio, i.e. the Bühlmann credibility estimator $\Pi_{i,n+1}^{B}$ of $\mu(\theta_i)$ is by definition the best estimator in the class

$$\Pi_{i,n+1}^{B} : \Pi_{i,n+1}^{B} = c_0^i + \sum_{j=1}^{I}\sum_{t=1}^{n} c_{jt}^i X_{jt}; \qquad c_0^i, c_{jt}^i \in \mathbb{R}.$$

Because of independence between contracts, we know that the credibility premium of contract i is a function of its observations only, we can write it in this form:

$$\Pi_{i,n+1}^{B} = c_0 + \sum_{t=1}^{n} c_t X_{it}.$$

Taking partial derivatives with respect to c_0 and c_t, $t = 1, ..., n$, we have:

$$\begin{aligned} c_0 &= E\left[\mu(\theta_i)\right] - \sum_{t=1}^{n} c_t E\left[X_{it}\right] = (1 - nc)\, m, \\ c_1 &= .... = c_n = c = \frac{a}{an + s^2} \quad , \end{aligned}$$

Then,

$$\begin{aligned} \Pi_{i,n+1}^{B} &= nc\bar{X}_i + (1 - nc)\, m, \\ &= z\bar{X}_i + (1 - z)\, m, \qquad (1.10) \end{aligned}$$

Where

$$
\begin{aligned}
z &= \frac{n}{n+k} = \frac{n}{n+\frac{s^2}{a}}, \\
\bar{X}_i &= \frac{1}{n}\sum_{t=1}^{n} X_{it},
\end{aligned}
$$

$$k = \frac{s^2}{a} \text{ : is called the credibility coefficient,}$$

$$
\begin{aligned}
s^2 &= E\left[\sigma^2\left(\theta\right)\right], \\
a &= Var\left[\mu(\theta)\right].
\end{aligned}
$$

Remark 1.3

The structure parameter s^2 is a global measure of the stability of the portfolio's claim experience. s^2 is sometimes called the "homogeneity within the insureds". The lower the value of s^2, the more stable the portfolio's claim experience, the larger the credibility factor.

The structure parameter a is a measure of the stability of the various individual risk premiums and is sometimes referred to as the "homogeneity between the insureds". In other words, a is an indicator of the heterogeneity of the portfolio's experience. The greater the heterogeneity of a portfolio, the more important is the weight given to individual experience. Hence, as a increases, z increases also.

1.3.2 Parametric approach (pure Bayesian approach)

Example 1.2 (Poisson-gamma)

Taking

$$f(x \mid \theta) = \frac{\theta^x e^{-\theta}}{x!}, x = 0, 1, ...$$

and

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \quad \theta > 0, \alpha > 0, \beta > 0.$$

We know that $\mu(\theta) = E[X \mid \theta] = \theta$ and $\sigma^2(\theta) = var[X \mid \theta] = \theta$, consequently,

$$m = E[\theta] = \frac{\alpha}{\beta},$$

$$s^2 = E[\theta] = \frac{\alpha}{\beta},$$

$$a = Var[\theta] = \frac{\alpha}{\beta^2},$$

$$k = \frac{s^2}{a} = \beta,$$

$$z = \frac{n}{n+\beta},$$

Finally,

$$\Pi_{i,n+1}^{B} = \frac{n}{n+\beta}\bar{X}_i + \left(1 - \frac{n}{n+\beta}\right)\frac{\alpha}{\beta}.$$

Example 1.3 (geometric-beta)

Suppose

$$f(x \mid \theta) = \theta(1-\theta)^x, x = 0, 1, ...$$

and

$$\pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, \quad 0 < \theta < 1, \alpha > 0, \beta > 0,$$

$$\mu(\theta) = \frac{1-\theta}{\theta},$$

$$\sigma^2(\theta) = \frac{1-\theta}{\theta^2},$$

$$m = \frac{\beta}{\alpha-1},$$

$$k = \frac{s^2}{a} = \alpha - 1,$$

$$z = \frac{n}{n+\alpha-1},$$

Then

$$\Pi_{i,n+1}^{B} = \frac{n}{n+\alpha-1}\bar{X}_i + \left(1 - \frac{n}{n+\alpha-1}\right)\frac{\beta}{\alpha-1}.$$

1.3.3 Non parametric approach (empirical Bayesian approach)

The parametric approach has a limited interest in practice, because its necessity to knowing $f(x \mid \theta)$ and $\pi(\theta)$. We assume, as many actuaries do in practical situations, that we know nothing about such probability distributions. These actuaries feel more comfortable in relying exclusively on the data, rather than making subjective judgments about the prior mean and other parameter values. These actuaries estimate the parameters of interest using the available data.

Using the non parametric approach, we replace the pure Bayesian approach by the empirical Bayesian approach.

Estimation of the structure parameters

The structure parameters m, s^2, and a are unknown in practice. Hence, they must be estimated from the entire portfolio data.

We employ the following estimators for m, s^2, and a, respectively:

$$\hat{m} = \bar{X}_{\bullet\bullet} = \frac{1}{I}\sum_{i=1}^{I}\bar{X}_{i\bullet} = \frac{1}{In}\sum_{i=1}^{I}\sum_{t=1}^{n}X_{it},$$

where

$$\bar{X}_{i\bullet} = \frac{1}{n}\sum_{t=1}^{n}X_{it},$$

$$\hat{s}^2 = \frac{1}{I}\sum_{i=1}^{I}\hat{s}_i^2 = \frac{1}{I(n-1)}\sum_{i=1}^{I}\sum_{t=1}^{n}\left(X_{it} - \bar{X}_{i\bullet}\right)^2,$$

where

$$\hat{s}_i^2 = \frac{1}{(n-1)} \sum_{t=1}^{n} \left(X_{it} - \bar{X}_{i\bullet}\right)^2,$$

and

$$\hat{a} = \frac{1}{I-1} \sum_{i=1}^{I} \left(\bar{X}_{i\bullet} - \bar{X}_{\bullet\bullet}\right)^2 - \frac{\hat{s}^2}{n},$$

$$= \frac{1}{I-1} \sum_{i=1}^{I} \left(\bar{X}_{i\bullet} - \bar{X}_{\bullet\bullet}\right)^2 - \frac{1}{In\,(n-1)} \sum_{i=1}^{I} \sum_{t=1}^{n} \left(X_{it} - \bar{X}_{i\bullet}\right)^2.$$

It turns out that $\hat{m}$, $\hat{s}^2$, and $\hat{a}$ are unbiased estimators of m, s^2, and a, respectively. The estimators of k and z are then $\hat{k} = \frac{\hat{s}^2}{\hat{a}}$ and $\hat{z} = \frac{n}{n+\hat{k}}$, where neither $\hat{k}$ nor $\hat{z}$ is an unbiased estimator. Finally, the Bühlmann estimate is:

$$\hat{\Pi}_{i,n+1}^{B} = \hat{z}\bar{X}_{i\bullet} + (1-\hat{z})\,\hat{m}, \quad \text{for } i = 1, 2, ..., I.$$

It is possible that $\hat{a}$ can be negative, an undesirable result since variances must, of course, be nonnegative. In practice, we suppose $\hat{a}^{'} = \max(\hat{a}, 0)$ which is a biased estimator. The non parametric approach is illustrated in the following examples.

Example 1.4

An insurance company has two group workers' compensation policies. The aggregate claim amounts in thousands of dinars for the first three policy years are summarized in the table below. Using Bühlmann's model to estimate the aggregate claim amount during the fourth policy year for each of the two group policies. We don't make any assumptions about the probability distribution of aggregate claim amounts.

	Aggregate claim amounts		
Group	Policy year		
policy	1	2	3
1	5	8	11
2	11	13	12

Table 1.3 - Aggregate Claim Amounts

Solution:

Since there are two group policies and three years of experience data for each, we have $I = 2$ and $n = 3$. The observed claim vectors are $X_1 = (X_{11}, X_{12}, X_{13}) = (5, 8, 11)$ and $X_2 = (X_{21}, X_{22}, X_{23}) = (11, 13, 12)$ respectively, for the two policies. Since:

$$\bar{X}_{1\bullet} = \frac{1}{n}\sum_{t=1}^{n} X_{1t} = \frac{5+8+11}{3} = 8,$$

and

$$\bar{X}_{2\bullet} = \frac{1}{n}\sum_{t=1}^{n} X_{2t} = \frac{11+13+12}{3} = 12,$$

Then, the estimate of the overall mean, $\hat{m}$, is

$$\hat{m} = \bar{X}_{\bullet\bullet} = \frac{1}{I}\sum_{i=1}^{I} \bar{X}_{i\bullet} = \frac{8+12}{2} = 10,$$

Furthermore, since

$$\hat{s}_1^2 = \frac{1}{(n-1)}\sum_{t=1}^{n} \left(X_{1j} - \bar{X}_{1\bullet}\right)^2 = \frac{1}{2}\left[(5-8)^2 + (8-8)^2 + (11-8)^2\right] = 9,$$

and

$$\hat{s}_2^2 = \frac{1}{(n-1)}\sum_{t=1}^{n} \left(X_{2j} - \bar{X}_{2\bullet}\right)^2 = \frac{1}{2}\left[(11-12)^2 + (13-12)^2 + (12-12)^2\right] = 1,$$

The estimate of the expected process variance is

$$\hat{s}^2 = \frac{1}{I}\sum_{i=1}^{I} \hat{s}_i^2 = \frac{1}{2}(9+1) = 5.$$

The estimate of the variance of the hypothetical means is

$$\hat{a} = \frac{1}{I-1}\sum_{i=1}^{I} \left(\bar{X}_{i\bullet} - \bar{X}_{\bullet\bullet}\right)^2 - \frac{\hat{s}^2}{n} = \frac{1}{1}\left[(8-10)^2 + (12-10)^2\right] - \frac{5}{3} = \frac{19}{3},$$

Then we have

$$\hat{k} = \frac{\hat{s}^2}{\hat{a}} = \frac{5}{\frac{19}{3}} = \frac{15}{19} = 0.78947,$$

Next, the estimated credibility factor for each group policy is

$$\hat{z} = \frac{n}{n + \hat{k}} = \frac{3}{3 + 0.78947} = 0.79167,$$

Finally, the estimated aggregate claim amounts for the fourth year are

$$\begin{aligned} \hat{\Pi}^B_{1,4} &= \hat{z}\bar{X}_{1\bullet} + (1 - \hat{z})\,\hat{m} = (0.79167)(8) + (0.20833)(10) = \ 8.41666, \\ \hat{\Pi}^B_{2,4} &= \hat{z}\bar{X}_{2\bullet} + (1 - \hat{z})\,\hat{m} = (0.79167)(12) + (0.20833)(10) = 11.58334, \end{aligned}$$

respectively.

Example 1.5

This example is from [25], which is based on [27]. These informations represent the amount claim in automobile insurance between July 1970 and 1973 in 5 American states. We have $I = 5$ contracts and $n = 12$ experience periods. The average claim amounts X_{it} are presented in the following table.

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
$t=1$	1738	1364	1759	1223	1456
$t=2$	1642	1408	1685	1146	1499
$t=3$	1794	1597	1479	1010	1609
$t=4$	2051	1444	1763	1257	1471
$t=5$	2079	1342	1674	1426	1482
$t=6$	2234	1675	2103	1532	1572
$t=7$	2032	1470	1502	1953	1606
$t=8$	2035	1448	1622	1123	1735
$t=9$	2115	1464	1828	1343	1607
$t=10$	2262	1831	2155	1243	1573
$t=11$	2267	1612	2233	1762	1613
$t=12$	2517	1471	2059	1306	1690

Table 1.4 - Average claim amounts (X_{it})

in the Hachmeister portfolio

The estimators of the structure parameters are:

$$\begin{aligned} \hat{m} &= 1671, \\ \hat{s}^2 &= 46040, \\ \hat{a} &= 72310, \end{aligned}$$

which are regrouped in the table below

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
Individual premium $\bar{X}_i$	2064	1511	1822	1360	1599
Credibility premium $\hat{\Pi}^B_{i,13}$	2044	1519	1814	1376	1602
Credibility factor z	0.95	0.95	0.95	0.95	0.95

Table 1.5 - Results with the Bühlmann model

Summary of the non parametric approach

Goal: To determine the compromise estimate of the aggregate claim amount for each of the I policyholders.

Step 1: Determine the number of policyholders, $I \geq 2$.

Step 2: Determine the number of policy years of experience, $n \geq 2$

Step 3: Compute the aggregate claim amount, X_{it}, for each policyholder during each policy year.

Step 4: Compute the average claim amount, $\bar{X}_{i.}$, over all policy years for each policyholder.

Step 5: Compute the estimated overall mean, $\hat{m}$.

Step 6: Compute the estimated expected process variance, $\hat{s}^2$.

Step 7: Compute the estimated variance of the hypothetical means, $\hat{a}$.

Step 8: Calculate $\hat{k} = \frac{\hat{s}^2}{\hat{a}}$.

Step 9: Compute the credibility factor, $\hat{z} = \frac{n}{n+\hat{k}}$.

Step 10: Compute the Bühlmann credibility estimate, $\hat{\Pi}^B_{i,n+1}$, of the aggregate claim amount for each policyholder.

Remark 1.4

Besides Bayesian and Bühlmann models, a number of other credibility models have also been introduced by various researchers like the Bühlmann-Straub model which generalized the Bühlmann model by associating a different credibility factors for each contract.

Moreover, we find models including the random coefficients regression credibility model introduced by [27], the hierarchical credibility model and crossed classification credibility model. For details and formulas see surveys by [7], [9], [10], [26] and [41].

Chapter 2

Credibility Premium Under Entropy Loss Function

Under the frequentist paradigm of statistics, every parameter, θ, is assumed to be a fixed but unknown quantity with an underlying "true value". Statistical inference involves constructing either:

- A point estimate of θ or
- A confidence interval around θ.

In contrast, under the Bayesian paradigm, every parameter is assumed to be a random variable. Before any data are observed, θ is assumed to have a particular prior distribution. After the observation of pertinent data, a revised or posterior distribution can be computed for θ via Bayes theorem.

A decision problem is generally based on 3 elements:

- Action (decision) space D.
- Set of parameters Θ.
- Loss function $L(\theta,\, \hat{\theta})$.

In this chapter, we begin by giving some basic elements about decision theory and loss functions. After, instead of traditional squared error loss function, we define and present the entropy loss function.

2.1 Basic elements

Let $\hat{\theta} \in D$ *a decision rule.*

A loss function $L(\theta,\, \hat{\theta})$ *is a measurable function* $(\Theta \times D)$ *with values in* $\mathbb{R}_+$ *which is defined as*

- $\forall(\theta,\, \hat{\theta}) : L(\theta,\, \hat{\theta}) > 0.$
- $\forall\theta, \exists\, \tilde{\theta} : L(\theta,\, \tilde{\theta}) = 0.$

In order to obtain a point estimate of θ from its posterior distribution, it is necessary to specify a loss function of θ, $L(\theta,\, \hat{\theta})$, where we use $\hat{\theta}$ to denote an estimator of θ.

We can write $L(\theta,\, \hat{\theta})$: loss, if θ is the "true" parameter and $\hat{\theta}$ is the value taken by the estimator when the data is observed.

As in Bayesian decision analysis (see, for example, [32]), we employ the estimator of θ that minimizes the expected value of $L(\theta, \hat{\theta})$.

From this, we derive the risk function of the estimator $\hat{\theta}$

$$R\left(\theta,\hat{\theta}\right) := E_\theta\left[L(\theta,\hat{\theta})\right] = \int_{R^n} L(\theta,\hat{\theta})dF_\theta\left(x\right). \tag{2.1}$$

The goal then is to find an estimator $\hat{\theta} \in D$, for which the risk function $R\left(\theta,\hat{\theta}\right)$ is as small as possible. In general, it is not possible to do this simultaneously for all values of θ i.e. there is no $\hat{\theta}$ which minimizes $R\left(\theta,\hat{\theta}\right)$ uniformly over θ. In *Figure* 2.1, we see an example, where depending on the value of θ, $\hat{\theta}_1$ or $\hat{\theta}_2$ has the smaller value of the risk function.

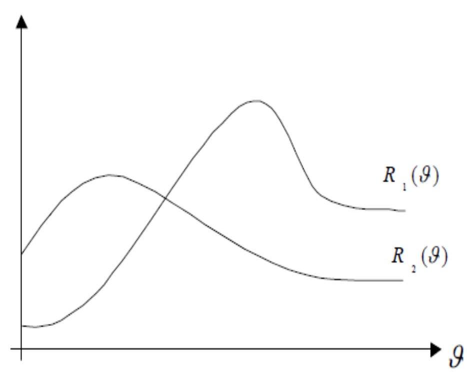

Figure 2.1 -Risk functions for $R_1\left(\theta\right)$ and $R_2\left(\theta\right)$

The estimator $\hat{\theta}$ has two main properties:

Admissibility:

$\hat{\theta}$ is admissible if there is no estimator $\tilde{\theta}$ such

$$R\left(\theta,\tilde{\theta}\right) \leq R\left(\theta,\hat{\theta}\right) \ \forall\theta \ \text{ and } \quad R\left(\theta_0,\tilde{\theta}\right) \leq R\left(\theta_0,\hat{\theta}\right).$$

If $\hat{\theta}$ is unique, then $\hat{\theta}$ is admissible.

Minimaxity:

$\hat{\theta}$ is minimax if

$$\forall\tilde{\theta}, \sup_{\theta} R\left(\theta,\hat{\theta}\right) \leq R\left(\theta,\tilde{\theta}\right),$$

i.e.

$$\sup_{\theta} R\left(\theta,\hat{\theta}\right) = \inf_{\tilde{\theta}} \sup_{\theta} R\left(\theta,\tilde{\theta}\right).$$

Remark 2.1

In actuarial sciences, when we studying the derivation of credibility premiums, as a statistical decision-making process, some Bayesian premium estimators P^B of the risk premium $\mu(\theta)$ are derived under the symmetric squared error loss function (SELF, henceforth), $L(\mu(\theta), \mathrm{P}^B) = \left(\mu(\theta) - \mathrm{P}^B\right)^2$,which is often used also because it does not lead to extensive numerical computation because the Bayesian premium P^B which minimizes the expected value of the loss function is the mean of the posterior distribution of θ, but the SELF considers a same penalty for a policyholder to be under or overcharged and gives no credit to situations in which the policyholder is undercharged and the insurance company might lose its money. There are numerous examples in the literature that suggest the use of loss function associated with estimation or prediction that should assign a more severe penalty for overestimation than underestimation or vice versa (see:[48]; [6]; and [36]). Hence, the use of the symmetric SELF is critical and it is beneficial to consider some loss functions -like entropy loss- which assigned less or more penalty for being under or overcharged.

2.2 The entropy loss function

In many practical situations, it appears to be more realistic to express the loss in terms of the ratio $\frac{\hat{\theta}}{\theta}$, in this case, [11] proposed a loss function which is called **general** entropy Loss function (which is also known as Stein loss), it is a useful asymmetric loss when giving credit to under and overestimation which has this form

$$L(\theta, \hat{\theta}) = \left(\frac{\hat{\theta}}{\theta}\right)^{q} - q \ln\left(\frac{\hat{\theta}}{\theta}\right) - 1, \qquad q \neq 0, \tag{2.2}$$

whose minimum occurs at $\hat{\theta} = \theta$,

q is the loss parameter which reflects the departure from symmetry. The loss parameter q allows different shapes of this loss function. For $q > 0$, a positive error (overestimation) has a more serious effect than a negative error, and for $q < 0$, a negative error (underestimation) has a more serious effect than a positive error.

$$E_\theta\left[L(\theta,\hat{\theta})\right]=\left(\hat{\theta}\right)^q E_\theta\left[\theta^{-q}|\underline{\mathrm{X}}\right]-q\ln\left(\hat{\theta}\right)+q\ln\left(\theta\right)-1$$

To achieve the minimum in the above equation, the derivative must be set to zero, namely

$$\frac{d}{d\theta}E_\theta\left[L(\theta,\hat{\theta})\right]=q\left(\hat{\theta}\right)^{q-1}E_\theta\left[\theta^{-q}|\underline{\mathrm{X}}\right]-\frac{q}{\hat{\theta}}=0,$$

$$\hat{\theta}^q=E_\theta\left[\theta^{-q}|\underline{\mathrm{X}}\right]^{-1},$$

$$\hat{\theta}=E_\theta\left[\theta^{-q}|\underline{\mathrm{X}}\right]^{-\frac{1}{q}}. \tag{2.3}$$

provided that the expectation $E_\theta\left[\theta^{-q}|\underline{\mathrm{X}}\right]$ exists and is finite.

As we know, the objective in the actuarial science is to find the Bayesian estimator of the individual premium , under the entropy loss, we have

$$L(\mu\left(\theta\right),\mathrm{P}^B_{\mathrm{ENT}})=\left(\frac{\mathrm{P}^B_{\mathrm{ENT}}}{\mu\left(\theta\right)}\right)^q-q\ln\left(\frac{\mathrm{P}^B_{\mathrm{ENT}}}{\mu\left(\theta\right)}\right)-1,\qquad q\neq 0,$$

Then

$$\mathrm{P}^B_{\mathrm{ENT}}=E_\theta\left[\mu\left(\theta\right)^{-q}|\underline{\mathrm{X}}\right]^{-q}. \tag{2.4}$$

In the rest of this work, we assume the original form where $q=1>0$, which has been used in [14] and [15]. Thus,

$$P^B_{\mathrm{ENT}} = E_\theta \left[\mu(\theta)^{-1} | \underset{\sim}{X} \right]^{-1}.$$

Note that for $q = -1$, the Bayesian premium under this loss, P^B_{ENT}, coincides with the Bayesian premium under the SELF.

$$P^B_{\mathrm{ENT}} = P^B_{\mathrm{SELF}}.$$

If we use the exponential family and the conjugate priors, we find that the Bayesian premium under the entropy loss function is a credibility premium (compromise between the prior mean and the mean of the current observations) only if $q = -1$.

Remark 2.2

The usual credibility formula holds whenever

- Claim size distribution is a member of the exponential family of distributions,
- Prior distribution conjugates with claim size distribution,
- Squared error loss has been considered. As long as, one of these conditions is violent, the usual credibility formula no longer holds.

If we change the squared error loss by the entropy loss function ($q \neq -1$) and taking the copula Poisson-gamma mentioned in example 1.1, we have

$$P^B_{\mathrm{ENT}} = E_\theta \left[\mu(\theta)^{-1} | \underset{\sim}{X} \right]^{-1} = \left[E_\theta \left[\frac{1}{\theta} | \underset{\sim}{X} \right] \right]^{-1} = \left[\frac{\beta + n}{\alpha + \sum_i x_i - 1} \right]^{-1} = \frac{\alpha + \sum_i x_i - 1}{\beta + n},$$

Thus

$$\mathrm{P}_{\mathrm{ENT}}^{B}=\frac{n}{n+\beta}\bar{X}\quad+\frac{\alpha-1}{n+\beta}.$$

- Bernoulli-Beta:

$$\mathrm{P}_{\mathrm{ENT}}^{B}=E_{\theta}\left[\mu\left(\theta\right)^{-1}|\underset{\sim}{\mathrm{X}}\right]^{-1}=E_{\theta}\left[\frac{1}{\theta}|\underset{\sim}{\mathrm{X}}\right]^{-1}=\left[\frac{\alpha+\beta+n-1}{\alpha+\sum_{i}x_{i}-1}\right]^{-1}=\frac{\alpha+\sum_{i}x_{i}-1}{\alpha+\beta+n-1},$$

Thus

$$\mathrm{P}_{\mathrm{ENT}}^{B}=\frac{n}{n+\alpha+\beta-1}\bar{X}\quad+\frac{\alpha-1}{n+\alpha+\beta-1}.$$

we remark that $\mathrm{P}_{\mathrm{ENT}}^{B}$ is not a linear combination of the observation mean and mean of prior, it is just an affine function of the observation mean $\bar{X}$, then, we have not a credibility formula.

Payandeh (2010) , using the mean square error minimization technique, develops a simple and practical approach to the credibility theory. Namely, he approximates the Bayes estimator with respect to a general loss function and general prior distribution by a convex combination of the observation mean and mean of prior, say, approximate credibility formula.

It is well known that Bayes estimator reflects properties of loss function and prior distribution (see Marchand and Payandeh, 2009). Therefore, it makes sense to consider a Bayes estimator, under an appropriate loss and prior distribution, as a suitable and acceptable estimator which reflects our concerns about an unknown parameter.

The following lemma considers a Bayes estimator as an appropriate estimator for parameter. Then, using the mean square error technique develops a new approach to approximate the Bayes estimator by the credibility formula.

Lemme 2.1 $(Payandeh\ 2010)$. *Suppose $X_1, X_2, ..., X_n$ given risk parameter are identical and independently distributed with $\mu(\theta) = E[X \mid \theta]$ for $i = 1, 2, ..., n$. Moreover, suppose that the risk parameter θ has a prior distribution π with mean m, i.e., $m = E[\theta]$, and P^B is a Bayes estimator with respect to loss function L and prior distribution π.*

Then, in the class of credibility premiums

$$P = \left\{P_z \text{ :where } P_z = z\bar{X} + (1-z)\,m, \text{ and } z \in [0,1]\right\}$$

an estimator P_{opt}, with

$$z_{opt} = \frac{E\left[\left(\bar{X}-m\right)\left(\mathrm{P}^B - m\right)\right]}{E\left[\left(\bar{X}-m\right)\right]^2}, \tag{2.5}$$

minimizes the mean squared error between P^B and P_z, i.e.,

$$\mathrm{P}_{opt} = \arg\min E\left[\left(\mathrm{P}^B - \mathrm{P}_z\right)\right]^2. \tag{2.6}$$

Proof. Mean square distance between two estimators P^B and P_z can be readily observed as

$$MSE\,(z) = E\left[\left(\mathrm{P}^B - \mathrm{P}_z\right)\right]^2$$

$$= E\left[\left(\mathrm{P}^B - z\bar{X} - (1-z)\,m\right)\right]^2.$$

Taking derivative with respect to z

$$\frac{\partial MSE\,(z)}{\partial z} = 2E\left[\left(\bar{X} - m\right)\right]^2 - 2zE\left[\left(\bar{X} - m\right)\left(\mathrm{P}^B - m\right)\right] = 0,$$

Along with the fact that second derivative of $MSE\,(z)$ with respect to z,

$$MSE''\,(z) = 2E\left[\bar{X} - m\right]^2,$$

is nonnegative lead to the desire result. ■

Two-folded expectations in nominator and denominator of z_{opt} given, the above can be simplified as

$$\begin{aligned} E\left[\left(\bar{X} - m\right)\left(\mathrm{P}^B - m\right)\right] &= E\left[cov\left(\bar{X}, \mathrm{P}^B \mid \theta\right)\right] + cov\left(E\left[\bar{X} \mid \theta\right], E\left[\mathrm{P}^B \mid \theta\right]\right) + \\ &\quad (m_0 - m)\left(m^{\mathrm{P}^B} - m\right). \end{aligned}$$

where

$$\begin{aligned} m_0 &= E\left[E\left[X_i \mid \theta\right]\right], \\ m^{\mathrm{P}^B} &= E\left[E\left[\mathrm{P}^B \mid \theta\right]\right] \text{ for } i = 1, 2, ..., n. \end{aligned}$$

Lemme 2.2 (*Payandeh* 2010). *In the existence of exact credibility premium, the approximate credibility premium, given by P_{opt}, coincides with the exact one.*

Proof. If the exact credibility premium holds, the Bayes estimator P^B can be written as $\mathrm{P}^B = z\bar{X} + (1 - z)\, m$. From this fact, one can observe that

$$\begin{aligned}
z_{opt} &= \frac{E\left[\left(\bar{X} - m\right)\left(\mathrm{P}^B - m\right)\right]}{E\left[\left(\bar{X} - m\right)\right]^2} \\
&= \frac{E\left[\left(\bar{X} - m\right)\left(z\bar{X} + (1 - z)\, m - m\right)\right]}{E\left[\left(\bar{X} - m\right)\right]^2} \\
&= \frac{E\left[\left(\bar{X} - m\right)\left(z\bar{X} - zm\right)\right]}{E\left[\left(\bar{X} - m\right)\right]^2} \\
&= z\frac{E\left[\left(\bar{X} - m\right)\left(\bar{X} - m\right)\right]}{E\left[\left(\bar{X} - m\right)\right]^2} \\
&= z.
\end{aligned}$$

■

Chapter 3

Bayesian Premium Estimators under Squared Error, Entropy and Linex Loss Functions: with Informative and Non Informative Priors

In this chapter, considering the Lindley distribution as conditional distribution of $X \mid \theta$, we focus on estimation of the Bayesian premium under three loss functions (squared error which is symmetric, linex and entropy which are asymmetrics), using non-informative and informative priors (the extension of Jeffreys and Inverted

Gamma priors) respectively. Because its difficulty and non linearity, we use a numerical approximation for computing the Bayesian premium.

A comparison was made through a Monte Carlo simulation study on the performance of these estimators according to the mean square error (MSE). Results are summarized in tables and followed by concluding remarks.

3.1 Around Lindley distribution

In Bayesian analysis, the unknown parameter θ is regarded as being the value of a random variable from a given probability distribution, with the knowledge of some information about the value of parameter prior to observing the data $x_1, x_2 ..., x_n$.

The Lindley distribution is one of the most popular distributions of waiting time and reliability theory. It was originally developed by [39] and some classical statistics properties are investigated by [23]. [37] considered Lindley distribution for reliability estimation using maximum likelihood and Bayesian approach. [51] introduced a discrete version of Lindley distribution known as discrete Poisson-Lindley distribution.

The distribution of zero-truncated Poisson-Lindley was introduced by [23] and [56] introduced the Negative Binomial distribution as an alternative to zero-truncated Poisson-Lindley distribution. [22] introduced a two parameter weighted Lindley distribution and pointed that Lindley distribution is particularly useful in modeling biological data from mortality studies. In addition, [57] introduced a new distribution, named gamma Lindley distribution, based on mixtures of gamma $(2, \theta)$ and one-parameter Lindley distributions which is useful in modeling lifedata. Recently, a study of the effect of some loss functions on Bayes Estimate and posterior risk for the Lindley distribution are made by [50].

Let $x_1, x_2, ..., x_n$ be independent and identically distributed from Lindley distribution with an unknown parameter θ. The probability density function is given by:

$$\begin{cases} f(x,\theta) = & \frac{\theta^2(1+x)e^{-\theta x}}{1+\theta}, \quad x,\theta > 0 \\ & 0, \qquad \text{otherwise.} \end{cases} \tag{3.1}$$

It can be written as a mixture of gamma distribution $(2, \theta)$ with exponential distribution (θ), where $p_1 = \frac{\theta}{\theta+1}$ and $p_2 = 1 - p_1$.

The corresponding cumulative distribution function (c.d.f.) is:

$$F(x) = 1 - \frac{\theta + 1 + \theta x}{\theta + 1} e^{-\theta x}, \quad x > 0, \theta > 0.$$

Many authors showed that (3.1) provides a better model for waiting times and survival times data than the exponential distribution.

However, due to the popularity of the exponential distribution, the Lindley distribution has been overlooked in the actuarial literature and in many applied areas.

The expectation and the variance of Lindley distribution are

$$\begin{aligned} \mu(\theta) &= E[X \mid \theta] = \frac{\theta + 2}{\theta(\theta + 1)}, \\ \sigma^2(\theta) &= var[X \mid \theta] = \frac{\theta^2 + 4\theta + 2}{\theta^2(\theta + 1)^2}. \end{aligned}$$

The likelihood function for a random sample $x_1, x_2..., x_n$ which is taken from Lindley distribution is:

$$L\left(x,\ \theta\right)=\left(\frac{\theta^2}{1+\theta}\right)^n \Pi_{i=1}^n\left(1+x_i\right)e^{-\theta\sum_{i=1}^n x_i},\ x,\theta>0 \tag{3.2}$$

The logarithm of likelihood function is

$$\ln L\left(x,\ \theta\right)=n\left(\ln\theta^2-\ln\left(1+\theta\right)\right)+\sum_{i=1}^n\ln\left(1+x_i\right)-\theta\sum_{i=1}^n x_i.$$

Maximizing for θ

$$\frac{d\ln L\left(x,\ \theta\right)}{d\theta}=0.$$

We get

$$\bar{x}\theta^2+\left(\bar{x}-1\right)\theta-2=0,\ \theta>0.$$

Solving for θ we have

$$\hat{\theta}=\frac{1-\bar{x}+\sqrt{\left(\bar{x}-1\right)^2+8\bar{x}}}{2\bar{x}}.$$

It can be written under the form $\frac{p(x)e^{-\theta x}}{q(\theta)}$ with

$$p\left(x\right)=1+x,$$

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

which means that it's a member of the exponential family,

It is well known that, for Bayesian premium estimators, the performance depends on the form of the prior distribution and the loss function assumed.

3.2 Derivation of Bayesian premiums

To obtain Bayesian premium estimators, we assume that θ is a real valued random variable with probability density function $\pi(\theta)$. Recall that the conditional distribution of $X \mid \theta$ is the Lindley distribution and the distribution of θ is assumed to be known in the present section.

$f(\theta \mid \underset{\sim}{X})$ is the posterior distribution of θ given the data. In this section we consider the estimation of the Bayesian premium $\mathrm{P}_{\bullet}^{B}$ based on the above mentioned priors and loss functions.

3.2.1 Bayesian premium estimators under squared error loss function

The squared error loss function was proposed by [38] and [10] to develop least squares theory. It is defined as

$$L(\hat{\theta}, \theta) = \left(\hat{\theta} - \theta\right)^2,$$

To obtain Bayesian premium estimators, we write

$$L(\mathrm{P}_{\mathrm{SELF}}^{B}, \mu(\theta)) = \left(\mathrm{P}_{\mathrm{SELF}}^{B} - \mu(\theta)\right)^2,$$

The Bayesian premium $\mathrm{P}_{\mathrm{SELF}}^{B}$ is the estimator of the individual premium $\mu(\theta)$, it is to be chosen such that the posterior expectation of the squared error loss function

$$\begin{aligned} E\left[L(\mathrm{P}_{\mathrm{SELF}}^{B}, \mu(\theta))\right] &= \int_0^\infty L(\mathrm{P}_{\mathrm{SELF}}^{B}, \mu(\theta)) f(\theta \mid \underset{\sim}{\mathrm{X}})\, d\theta \\ &= \int_0^\infty \left(\mathrm{P}_{\mathrm{SELF}}^{B} - \mu(\theta)\right) f(\theta \mid \underset{\sim}{\mathrm{X}})\, d\theta, \end{aligned}$$

is minimum.

$$\begin{aligned} \text{Here} \qquad \mathrm{P}_{\mathrm{SELF}}^{B} &= E\left[\mu(\theta) \mid \underset{\sim}{\mathrm{X}}\right]. && (3.3) \\ &= \int_0^\infty \mu(\theta) f(\theta \mid \underset{\sim}{\mathrm{X}})\, d\theta, && ((3.4)) \end{aligned}$$

where

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

is the individual premium.

Posterior distribution using the extension of Jeffreys prior

Bayesian approach makes use of ones prior knowledge about the parameters as well as the available data. When ones prior knowledge about the parameter is not available or very little, it is possible to make use of the noninformative prior in Bayesian analysis.

Since we have no knowledge on the parameters, we seek to use the extension of Jeffreys' prior information, where Jeffreys' prior is the square root of the determinant of the Fisher information.

We find Jeffrey prior by taking $\pi(\theta) = \sqrt{I(\theta)}$, where

$$I(\theta) = -E\left[\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right] = \frac{\theta^2 + 4\theta + 2}{\theta^2 (1+\theta)^2},$$

is the Fisher information.

The extension of Jeffreys distribution is assumed as non-informative prior for the parameter θ. It was proposed by [2] and [30], it is given as:

$$\pi(\theta) = [I(\theta)]^c = k\left[\frac{\theta^2 + 4\theta + 2}{\theta^2 (1+\theta)^2}\right]^c, \ \theta, c > 0, \ k \text{ is a constant.} \tag{3.5}$$

Combining (3.5) with the likelihood function of the Lindley distribution, the posterior distribution of parameter θ given the data $(X_1, X_2 ..., X_n)$ is derived as follows:

$$
\begin{aligned}
f(\theta \mid \underset{\sim}{X}) &= \frac{\Pi_{i=1}^{n} L(x_i\ ,\theta)\,\pi(\theta)}{\int_0^\infty \Pi_{i=1}^{n} L(x_i\ ,\theta)\,\pi(\theta)\,d\theta} \\
&= \frac{\frac{\theta^{2(n-c)}}{(1+\theta)^{n+2c}}\left(\theta^2+4\theta+2\right)^c e^{-\theta\sum_{i=1}^{n} x_i}}{\int_0^\infty \frac{\theta^{2(n-c)}}{(1+\theta)^{n+2c}}\left(\theta^2+4\theta+2\right)^c e^{-\theta\sum_{i=1}^{n} x_i} d\theta}, \theta > 0
\end{aligned}
$$

According to the squared error loss function, the corresponding Bayesian premium estimator is derived by substituting the posterior distribution (3.5) in (3.3), as follows:

$$
\mathrm{P}_{\mathrm{SELF}}^{B} = \int_0^\infty \mu(\theta)\, f(\theta \mid \underset{\sim}{X})\, d\theta
$$

$$
= \frac{\int_0^\infty \frac{\theta^{2(n-c)-1}}{(1+\theta)^{n+2c+1}}\left(\theta^2+4\theta+2\right)^c (\theta+2)\, e^{-\theta\sum_{i=1}^{n} x_i} d\theta}{\int_0^\infty \frac{\theta^{2(n-c)}}{(1+\theta)^{n+2c}}\left(\theta^2+4\theta+2\right)^c e^{-\theta\sum_{i=1}^{n} x_i} d\theta}, \quad \theta > 0 \tag{3.6}
$$

We know that only combinations of unidimensional exponential family members with their natural conjugate priors yield linear Bayesian premiums (exact credibility formula) which are mentioned in the Table (1.1).

It may be noted here that the posterior distribution $f(\theta \mid \underset{\sim}{X})$ takes a ratio form that it give not a credibility formula and involves an integration in the denominator and cannot be reduced to a closed form. Hence, the evaluation of the posterior expectation for obtaining the Bayesian premium of θ will be tedious. Among the various methods suggested to approximate the ratio of integrals of the above form, perhaps the simplest one is Lindley's (1980) approximation method, which approaches the ratio of the integrals as a whole and produces a single numerical result. Thus, we propose the use of Lindley's (1980) approximation for obtaining the Bayesian premium. Many authors have used this approximation for obtaining the Bayes estimators for some distributions, see among others, [28] and [31].

If n is sufficiently large, according to Lindley (1980), any ratio of the integral of the form

$$I(x) = E[h(\theta)]$$

$$= \frac{\int_\theta h(\theta)\exp[L(\theta,x)+g(\theta)]\,d\theta}{\int_\theta \exp[L(\theta,x)+g(\theta)]\,d\theta}, \qquad \theta > 0 \tag{3.7}$$

where $h(\theta) =$ function of θ only,

$L(\theta, x) =$ log of likelihood,

$g(\theta)=$ log of prior of θ. Thus,

$$I(x) = h\left(\hat{\theta}\right) + 0.5\left[\left(\hat{h}_{\theta\theta} + 2\hat{h}_\theta\hat{p}_\theta\right)\hat{\sigma}_{\theta\theta}\right] + 0.5\left[\left(\hat{h}_\theta\hat{\sigma}_{\theta\theta}\right)\left(\hat{L}_{\theta\theta\theta}\hat{\sigma}_{\theta\theta}\right)\right] \tag{3.8}$$

where $\hat{\theta} = \text{mle of } \theta = \dfrac{-(\bar{x}-1)+\sqrt{(\bar{x}-1)^2+8\bar{x}}}{2\bar{x}}, \ \bar{x} > 0.$

$$\hat{h}_\theta = \frac{\partial h\left(\hat{\theta}\right)}{\partial\hat{\theta}},$$

$$\hat{h}_{\theta\theta} = \frac{\partial^2 h\left(\hat{\theta}\right)}{\partial\hat{\theta}^2},$$

$$\hat{p}_{\theta} = \frac{\partial \ g\left(\hat{\theta}\right)}{\partial \hat{\theta}}, \tag{3.9}$$

$$\hat{L}_{\theta\theta} = \frac{\partial^2 L\left(\hat{\theta}\right)}{\partial \hat{\theta}^2},$$

$$\hat{\sigma}_{\theta\theta} = -\frac{1}{\hat{L}_{\theta\theta}},$$

$$\hat{L}_{\theta\theta\theta} = \frac{\partial^3 L\left(\hat{\theta}\right)}{\partial \hat{\theta}^3}.$$

After substituting the value of $f\left(\theta \mid \underset{\sim}{X}\right)$, it may be written as:

$$\mathrm{P}_{\mathrm{SELF}}^{B} = E\left[\mu\left(\theta\right) \mid \underset{\sim}{X}\right] = \frac{\int_{\theta} \mu\left(\theta\right) \exp\left[L\left(\theta, x\right) + g\left(\theta\right)\right] d\theta}{\int_{\theta} \exp\left[L\left(\theta, x\right) + g\left(\theta\right)\right] d\theta}, \qquad \theta > 0 \tag{3.10}$$

Where

$$h\left(\theta\right) = \mu\left(\theta\right) = \frac{\theta + 2}{\theta\left(\theta + 1\right)},$$

$$L\left(\theta, x\right) = 2n \ln \theta - n \ln\left(1 + \theta\right) - \theta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \ln\left(1 + x_i\right) \ ,$$

$$g\left(\theta\right) = c\left[\ln\left(\theta^2 + 4\theta + 2\right) - 2\ln\left(\theta^2 + \theta\right)\right],$$

It may easily be verified that

$$\hat{h}_\theta = -\frac{\theta^2+4\theta+2}{\left(\theta^2+\theta\right)^2},$$

$$\hat{h}_{\theta\theta} = \frac{2\theta^4+14\theta^3+24\theta^2+16\theta+4}{\theta^3\left(\theta+1\right)^4},$$

$$\hat{p}_\theta = 2c\left(\frac{\theta+2}{\theta^2+4\theta+2} - \frac{2\theta+1}{\theta^2+\theta}\right),$$

$$\hat{L}_{\theta\theta} = -\frac{2n}{\theta^2} + \frac{n}{\left(1+\theta\right)^2},$$

$$\hat{\sigma}_{\theta\theta} = \frac{\theta^2\left(1+\theta\right)^2}{2n\left(1+\theta\right)^2 - n\theta^2},$$

$$\hat{L}_{\theta\theta\theta} = \frac{4n}{\theta^3} - \frac{2n}{\left(1+\theta\right)^3},$$

Then, we get

$$\begin{aligned} \mathrm{P}^{B}_{\mathrm{SELF}} &= E\left[\mu\left(\theta\right) \mid \underset{\sim}{\mathrm{X}}\right] \\ &= \frac{\hat{\theta}+2}{\hat{\theta}\left(\hat{\theta}+1\right)} + \left[\left[\begin{array}{c}\left(\frac{\hat{\theta}^4+7\hat{\theta}^3+12\hat{\theta}^2+8\hat{\theta}+2}{\hat{\theta}^3\left(\hat{\theta}+1\right)^4}\right)+ \\ 2c\left(\frac{\hat{\theta}+2}{\hat{\theta}^2+4\hat{\theta}+2} - \frac{2\hat{\theta}+1}{\hat{\theta}^2+\hat{\theta}}\right)\left(\frac{\hat{\theta}^2+4\hat{\theta}+2}{\left(\hat{\theta}^2+\hat{\theta}\right)^2}\right)\end{array}\right]\frac{\hat{\theta}^2\left(1+\hat{\theta}\right)^2}{2n\left(1+\hat{\theta}\right)^2 - n\hat{\theta}^2}\right] + \\ &\quad \left[\left[\frac{\hat{\theta}^2\left(1+\hat{\theta}\right)^2}{2n\left(1+\hat{\theta}\right)^2 - n\hat{\theta}^2}\right]^2\left[\frac{n}{\left(1+\hat{\theta}\right)^3} - \frac{2n}{\hat{\theta}^3}\right]\left[\frac{\hat{\theta}^2+4\hat{\theta}+2}{\left(\hat{\theta}^2+\hat{\theta}\right)^2}\right]\right]. \end{aligned} \tag{3.11}$$

Posterior distribution using the Inverted Gamma prior (IG)

Informative priors are those that deliberately insert information that actuaries have at hand. This seems like a reasonable and reasoned approach since previous scientific knowledge should play a role in statistical inference. An informative prior provides more information than the non-informative priors, therefore the analysis using these prior more accurate and informative than classical approach.

The inverted gamma prior is a good distribution which represents the reciprocal of a variable distributed according to the gamma distribution. It is observed that if θ has an inverted gamma (α, β) distribution, then $\frac{1}{\theta}$ has a gamma (α, β) distribution.

It is given as

$$\pi(\theta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \frac{1}{\theta^{\alpha+1}} e^{-\frac{\beta}{\theta}}; \alpha, \beta, \theta > 0 \tag{3.12}$$

The first two moments of $IG(\alpha, \beta)$ are

$$\begin{aligned} E(\theta) &= \frac{\beta}{\alpha - 1}, \\ E\left(\theta^2\right) &= \frac{\beta^2}{(\alpha - 1)(\alpha - 2)}, \end{aligned}$$

Thus

$$Var(\theta) = E\left(\theta^2\right) - E(\theta)^2 = \frac{\beta^2}{(\alpha - 1)^2 (\alpha - 2)},$$

Now, using the likelihood of Lindley distribution and the Inverted Gamma prior (IG), the posterior distribution for the parameter θ given the data $(x_1, x_2 ..., x_n)$ takes the form

$$
\begin{aligned}
f\left(\theta \mid \underset{\sim}{X}\right) &= \frac{\Pi_{i=1}^{n} L\left(x_i\ ,\theta\right)\pi\left(\theta\right)}{\int_0^\infty \Pi_{i=1}^{n} L\left(x_i\ ,\theta\right)\pi\left(\theta\right)d\theta} \\
&= \frac{\frac{\theta^{2n-(\alpha+1)}}{(1+\theta)^n} e^{-\frac{\beta}{\theta}-\theta\sum_{i=1}^{n} x_i}}{\int_0^\infty \frac{\theta^{2n-(\alpha+1)}}{(1+\theta)^n} e^{-\frac{\beta}{\theta}-\theta\sum_{i=1}^{n} x_i} d\theta}. \qquad (3.13)
\end{aligned}
$$

Now, according to the squared error loss function, the corresponding Bayes' estimator for the parameter θ is as follows:

$$
\mathrm{P}_{\mathrm{SELF}}^{B} = E\left[\mu\left(\theta\right) \mid \underset{\sim}{X}\right] = \int_0^\infty \mu\left(\theta\right) f\left(\theta \mid \underset{\sim}{X}\right) d\theta
$$

$$
= \frac{\int_0^\infty \frac{\theta^{2n-(\alpha+2)}}{(1+\theta)^{n+1}}\left(\theta+2\right) e^{-\frac{\beta}{\theta}-\theta\sum_{i=1}^{n} x_i} d\theta}{\int_0^\infty \frac{\theta^{2n-(\alpha+1)}}{(1+\theta)^n} e^{-\frac{\beta}{\theta}-\theta\sum_{i=1}^{n} x_i} d\theta}, \quad \theta > 0
$$

Following the procedure as discussed above, we have

$$
g\left(\theta\right) = \beta \ln \alpha - \ln \Gamma\left(\beta\right) - \left(\alpha+1\right)\ln\theta - \frac{\beta}{\theta},
$$

$$
\hat{h}_\theta = -\frac{\theta^2+4\theta+2}{\left(\theta^2+\theta\right)^2},
$$

$$
\hat{h}_{\theta\theta} = \frac{2\theta^4+14\theta^3+24\theta^2+16\theta+4}{\theta^3\left(\theta+1\right)^4},
$$

$$
\hat{p}_\theta = \frac{\beta}{\theta^2} - \frac{\alpha+1}{\theta},
$$

$$\hat{L}_{\theta\theta} = -\frac{2n}{\theta^2} + \frac{n}{(1+\theta)^2},$$

$$\hat{\sigma}_{\theta\theta} = \frac{\theta^2 (1+\theta)^2}{2n (1+\theta)^2 - n\theta^2},$$

$$\hat{L}_{\theta\theta\theta} = \frac{4n}{\theta^3} - \frac{2n}{(1+\theta)^3},$$

We get after simplification

$$\begin{aligned} \mathrm{P}^{B}_{\mathrm{SELF}} &= E\left[\mu(\theta) \mid \underline{\mathrm{X}}\right] \\ &= \frac{\hat{\theta}+2}{\hat{\theta}\left(\hat{\theta}+1\right)} + \left[\left[\begin{array}{c} \left(\frac{\hat{\theta}^4+7\hat{\theta}^3+12\hat{\theta}^2+8\hat{\theta}+2}{\hat{\theta}^3\left(\hat{\theta}+1\right)^4}\right) - \\ \left(\frac{\hat{\theta}^2+4\hat{\theta}+2}{\left(\hat{\theta}^2+\hat{\theta}\right)^2}\right)\left(\frac{\beta}{\hat{\theta}^2} - \frac{\alpha+1}{\hat{\theta}}\right) \end{array}\right]\left[\frac{\hat{\theta}^2\left(1+\hat{\theta}\right)^2}{2n\left(1+\hat{\theta}\right)^2 - n\,\hat{\theta}^2}\right]\right] + \\ &\quad \left[\left[\frac{\hat{\theta}^2\left(1+\hat{\theta}\right)^2}{2n\left(1+\hat{\theta}\right)^2 - n\hat{\theta}^2}\right]^2\left[\frac{n}{\left(1+\hat{\theta}\right)^3} - \frac{2n}{\hat{\theta}^3}\right]\left[\frac{\hat{\theta}^2+4\hat{\theta}+2}{\left(\hat{\theta}^2+\hat{\theta}\right)^2}\right]\right]. \end{aligned} \tag{3.14}$$

3.2.2 Bayesian premium estimators under Linex loss function

The linex (linear-exponential) loss function is an asymmetric loss function -the name linex is justified by the fact that this loss function rises approximately linearly on one side of zero and approximately exponentially on the other side-, it was introduced by [54], [49],[4], [47], [52] and [44].

It may be expressed as :

$$L(\hat{\theta},\theta) = \exp\left(a\left(\hat{\theta}-\theta\right)\right) - a\left(\hat{\theta}-\theta\right) - 1, \qquad a \neq 0 \tag{3.15}$$

The sign and magnitude of the shape parameter a reflects the direction and degree of asymmetry, respectively. If $a > 0$, the overestimation is more serious than underestimation, and vice-versa. For a closed to zero, the linex loss is approximately squared error loss and therefore almost symmetric.

The posterior expectation of the linex loss function equation is :

$$E\left[L(\hat{\theta},\theta)\right] \propto \exp(a\hat{\theta})E\left[\exp(-a\theta)\right] - a\left(\hat{\theta} - E(\theta)\right) - 1,$$

By result of [58], the estimator of θ under the linex loss which minimizes the above equation is given by

$$\hat{\theta} = -\frac{1}{a}\ln\left[E\left[e^{-a\theta} \mid \underset{\sim}{X}\right]\right].$$

Taking derivatives to the posterior expectation with respect to $\hat{\theta}$, we have

$$a\exp(a\hat{\theta})E_\theta\left[\exp(-a\theta) \mid \underset{\sim}{X}\right] - a = 0,$$

$$\exp(-a\hat{\theta}) = E_\theta\left[\exp(-a\theta) \mid \underset{\sim}{X}\right],$$

$$\ln\exp(-a\hat{\theta}) = \ln E_\theta\left[\exp(-a\theta) \mid \underset{\sim}{X}\right],$$

$$\hat{\theta} = -\frac{1}{a} \ln \left[E_\theta \left[e^{-a\theta} \mid \underset{\sim}{X} \right] \right].$$

In our study, the aim is to find the Bayesian premium estimator $\mathrm{P}^B_{\mathrm{LIN}}$ which is the estimator of $\mu(\theta)$, then

$$\mathrm{P}^B_{\mathrm{LIN}} = -\frac{1}{a} \ln \left[E_\theta \left[e^{-a\mu(\theta)} \mid \underset{\sim}{X} \right] \right], \tag{3.16}$$

When the expectation $E_\theta \left[e^{-a\mu(\theta)} \right]$ exists and finite (see [11]).

Posterior distribution using the extension of Jeffreys prior

Using the linex loss function, the corresponding Bayes estimator of the parameter θ is as follows:

$$\mathrm{P}^B_{\mathrm{LIN}} = -\frac{1}{a} \ln E \left[e^{-a\mu(\theta)} \mid \underset{\sim}{X} \right].$$

$$\begin{aligned} E \left[e^{-a\mu(\theta)} \mid \underset{\sim}{X} \right] &= \int_0^\infty e^{-a\,\mu(\theta)} f\left(\theta \mid \underset{\sim}{X}\right) d\theta \\ &= \frac{\int_0^\infty \frac{\theta^{2(n-c)}}{(1+\theta)^{n+2c}} \left(\theta^2 + 4\theta + 2\right)^c e^{-\left(\theta \sum_{i=1}^n x_i + a\,\mu(\theta)\right)} d\theta}{\int_0^\infty \frac{\theta^{2(n-c)}}{(1+\theta)^{n+2c}} \left(\theta^2 + 4\theta + 2\right)^c e^{-\theta \sum_{i=1}^n x_i} d\theta} \\ &= \frac{\int_\theta h(\theta) \exp\left[L(\theta, x) + g(\theta)\right] d\theta}{\int_\theta \exp\left[L(\theta, x) + g(\theta)\right] d\theta}, \qquad \theta > 0 \end{aligned} \tag{3.17}$$

Following the same steps explained above, we have

$$h(\theta) = e^{-a\,\mu(\theta)},$$

$L(\theta, x)$ and $g(\theta)$ are the same as those given above,

$$\hat{h}_\theta = -a\mu^{'}(\theta)\, e^{-a\,\mu(\theta)},$$

$$\hat{h}_{\theta\theta} = -a\left(\mu^{''}(\theta) - a\mu^{'2}(\theta)\right) e^{-a\,\mu(\theta)},$$

$$\hat{p}_\theta = 2c\left(\frac{\theta+2}{\theta^2+4\theta+2} - \frac{2\theta+1}{\theta^2+\theta}\right),$$

$$\hat{L}_{\theta\theta} = -\frac{2n}{\theta^2} + \frac{n}{(1+\theta)^2},$$

$$\hat{\sigma}_{\theta\theta} = \frac{\theta^2(1+\theta)^2}{2n(1+\theta)^2 - n\theta^2},$$

$$\hat{L}_{\theta\theta\theta} = \frac{4n}{\theta^3} - \frac{2n}{(1+\theta)^3},$$

Where

$\mu(\theta)$ is the individual premium,

$$\mu^{'}(\theta) = \frac{\partial\mu(\theta)}{\partial\theta} = \left(\frac{\theta+2}{\theta(\theta+1)}\right)^{'} = -\frac{\theta^2+4\theta+2}{\left(\theta^2+\theta\right)^2},$$

$$\begin{aligned} \mu^{'}{}' (\theta) &= \frac{\partial^2 \mu(\theta)}{\partial \theta^2} = \left(\frac{\theta + 2}{\theta(\theta + 1)} \right)^{''} \\ &= \frac{2\theta^4 + 14\theta^3 + 24\theta^2 + 16\theta + 4}{\theta^3 (\theta + 1)^4}. \end{aligned}$$

$$\begin{aligned} E\left[e^{-a\mu(\theta)} \mid \underline{X}\right] &= e^{-a\,\mu\left(\hat{\theta}\right)} + \qquad (3.18) \\ & \left[\left(\begin{array}{c} \left(\left(a^2 A_1^2 - aA_2\right) e^{-a\,\mu\left(\hat{\theta}\right)} \right) + \\ \left(-4acA_1 e^{-a\,\mu\left(\hat{\theta}\right)} \right) A_3 \end{array} \right) \frac{\hat{\theta}^2 \left(1 + \hat{\theta}\right)^2}{4n\left(1 + \hat{\theta}\right)^2 - 2n\hat{\theta}^2} \right] + A_4. \end{aligned}$$

With

$$A_1 = \mu^{'}\left(\hat{\theta}\right) = \frac{\partial \mu(\theta)}{\partial \theta} = -\frac{\theta^2 + 4\theta + 2}{\left(\theta^2 + \theta\right)^2},$$

$$A_2 = \mu^{'}{}'(\theta) = \frac{\partial^2 \mu(\theta)}{\partial \theta^2} = \frac{2\theta^4 + 14\theta^3 + 24\theta^2 + 16\theta + 4}{\theta^3(\theta + 1)^4},$$

$$A_3 = \frac{\hat{\theta} + 2}{\hat{\theta}^2 + 4\hat{\theta} + 2} - \frac{2\hat{\theta} + 1}{\hat{\theta}^2 + \hat{\theta}}.$$

$$A_4 = \left[\left[\frac{\hat{\theta}^2 \left(1 + \hat{\theta}\right)^2}{2n\left(1 + \hat{\theta}\right)^2 - n\hat{\theta}^2} \right]^2 \left[aA_1 e^{-a\,\mu\left(\hat{\theta}\right)}\right] \left[\frac{n}{\left(1 + \hat{\theta}\right)^3} - \frac{2n}{\hat{\theta}^3} \right] \right].$$

Finally,

$$\mathrm{P}^{B}_{\mathrm{LIN}} = -\frac{1}{a} \ln E\left[e^{-a\mu(\theta)} \mid \underline{X}\right].$$

Posterior distribution using the Inverted Gamma prior (IG)

The corresponding Bayesian premium estimator under the linex loss function is:

$$\mathrm{P}_{\mathrm{LIN}}^{B} = -\frac{1}{a} \ln E\left[e^{-a\mu(\theta)} \mid \underset{\sim}{X}\right] \ .$$

$$\begin{aligned} E\left[e^{-a\mu(\theta)} \mid \underset{\sim}{X}\right] &= \int_0^\infty e^{-a\mu(\theta)} f\left(\theta \mid \underset{\sim}{X}\right) d\theta \\ &= \frac{\int_0^\infty \frac{\theta^{2n-(\alpha+1)}}{(1+\theta)^n} e^{-\left(\frac{\beta}{\theta}+\theta\sum_{i=1}^n x_i + a\mu(\theta)\right)} d\theta}{\int_0^\infty \frac{\theta^{2n-(\alpha+1)}}{(1+\theta)^n} e^{-\left(\frac{\beta}{\theta}+\theta\sum_{i=1}^n x_i\right)} d\theta}. \end{aligned} \tag{3.19}$$

Thus,

$$\hat{h}_\theta = -a\mu^{'}\left(\theta\right) e^{-a\ \mu(\theta)},$$

$$\hat{h}_{\theta\theta} = -a\left(\mu^{''}\left(\theta\right) - a\left(\mu^{'}\left(\theta\right)\right)^2\right) e^{-a\ \mu(\theta)},$$

$$\hat{p}_\theta = \frac{\beta}{\theta^2} - \frac{\alpha+1}{\theta},$$

$$\hat{L}_{\theta\theta} = -\frac{2n}{\theta^2} + \frac{n}{\left(1+\theta\right)^2},$$

$$\hat{\sigma}_{\theta\theta} = \frac{\theta^2\left(1+\theta\right)^2}{2n\left(1+\theta\right)^2 - n\theta^2},$$

$$\hat{L}_{\theta\theta\theta} = \frac{4n}{\theta^3} - \frac{2n}{(1+\theta)^3},$$

$$E\left[e^{-a\mu(\theta)} \mid \underset{\sim}{X}\right] = e^{-a\,\mu(\hat{\theta})} + A_5 \left[\begin{array}{c} 0.5\left(\left(a^2 A_1^2 - aA_2\right) e^{-a\,\mu(\hat{\theta})}\right) + \\ \left(-aA_1 e^{-a\,\mu(\hat{\theta})}\right)\left(\frac{\beta}{\hat{\theta}^2} - \frac{\alpha+1}{\hat{\theta}}\right) \end{array}\right] + A_4. \tag{3.20}$$

With

$$A_5 = \frac{\hat{\theta}^2\left(1+\hat{\theta}\right)^2}{2n\left(1+\hat{\theta}\right)^2 - n\hat{\theta}^2}.$$

Then,

$$\mathrm{P}^B_{\mathrm{LIN}} = -\frac{1}{a}\ln E_\theta\left[e^{-a\mu(\theta)} \mid \underset{\sim}{X}\right].$$

3.2.3 Bayesian premium estimators under entropy loss function

Posterior distribution using The extension of Jeffreys prior

Using the entropy loss function, the corresponding Bayesian premium estimator is as follows

$$\mathrm{P}^B_{\mathrm{ENT}} = \left(E\left[\mu(\theta)^{-1} \mid \underset{\sim}{X}\right]\right)^{-1}.$$

$$\begin{aligned} E\left[\mu(\theta)^{-1} \mid \underset{\sim}{X}\right] &= \int_0^{\infty} \mu(\theta)^{-1} f(\theta \mid \underset{\sim}{X})\, d\theta \\ &= \frac{\int_0^{\infty} \frac{\theta^{2(n-c)+1}}{(1+\theta)^{n+2c-1}} \frac{\left(\theta^2+4\theta+2\right)^c}{\theta+2} e^{-\theta \sum_{i=1}^n x_i} d\theta}{\int_0^{\infty} \frac{\theta^{2(n-c)}}{(1+\theta)^{n+2c}} \left(\theta^2+4\theta+2\right)^c e^{-\theta \sum_{i=1}^n x_i} d\theta} \\ &= \frac{\int_\theta h(\theta) \exp\left[L(\theta, x) + g(\theta)\right] d\theta}{\int_\theta \exp\left[L(\theta, x) + g(\theta)\right] d\theta}, \qquad \theta > 0 \end{aligned} \tag{3.21}$$

$$h(\theta) = \mu(\theta)^{-1} = \frac{1}{\mu(\theta)} = \frac{\theta+2}{\theta(\theta+1)},$$

$$\hat{h}_\theta = \frac{\theta^2+4\theta+2}{(\theta+2)^2},$$

$$\hat{h}_{\theta\theta} = \frac{4}{(\theta+2)^3},$$

$$\hat{p}_\theta = 2c\left(\frac{\theta+2}{\theta^2+4\theta+2} - \frac{2\theta+1}{\theta^2+\theta}\right),$$

$$\hat{L}_{\theta\theta} = -\frac{2n}{\theta^2} + \frac{n}{(1+\theta)^2},$$

$$\hat{\sigma}_{\theta\theta} = \frac{\theta^2(1+\theta)^2}{2n(1+\theta)^2 - n\theta^2},$$

$$\hat{L}_{\theta\theta\theta} = \frac{4n}{\theta^3} - \frac{2n}{(1+\theta)^3},$$

$$\mathrm{P}_{\mathrm{ENT}}^{B}=\left[\frac{\frac{\hat{\theta}+2}{\hat{\theta}\left(\hat{\theta}+1\right)}+\left[\left[\frac{2}{\left(\hat{\theta}+2\right)^{3}}+\frac{\hat{\theta}^{2}+2\hat{\theta}+2}{\left(\hat{\theta}+2\right)^{2}}\right]\left[\frac{\hat{\theta}+2}{\hat{\theta}^{2}+4\hat{\theta}+2}-\frac{2\hat{\theta}+1}{\hat{\theta}^{2}+\hat{\theta}}\right]\left[\frac{2c\hat{\theta}^{2}\left(1+\hat{\theta}\right)^{2}}{2n\left(1+\hat{\theta}\right)^{2}-n\hat{\theta}^{2}}\right]\right]+}{\left[\left[\frac{\hat{\theta}^{2}\left(1+\hat{\theta}\right)^{2}}{2n\left(1+\hat{\theta}\right)^{2}-n\hat{\theta}^{2}}\right]^{2}\left[\frac{\hat{\theta}^{2}+4\hat{\theta}+2}{\left(\hat{\theta}+2\right)^{2}}\right]\left[\frac{2n}{\hat{\theta}^{3}}-\frac{n}{\left(1+\hat{\theta}\right)^{3}}\right]\right]}\right]^{-1}. \tag{3.22}$$

Posterior distribution using the Inverted Gamma prior (IG)

We have

$$E\left[\mu\left(\theta\right)^{-1}\mid \underset{\sim}{X}\right]=\int_{0}^{\infty}\mu\left(\theta\right)^{-1}f\left(\theta\mid \underset{\sim}{X}\right)d\theta=\frac{\int_{0}^{\infty}\frac{\theta^{2n-\alpha}}{(1+\theta)^{n-1}(\theta+2)}e^{-\frac{\beta}{\theta}-\theta\sum_{i=1}^{n}x_{i}}d\theta}{\int_{0}^{\infty}\frac{\theta^{2n-(\alpha+1)}}{(1+\theta)^{n}}e^{-\frac{\beta}{\theta}-\theta\sum_{i=1}^{n}x_{i}}d\theta}.$$

$$\hat{h}_{\theta}=\frac{\theta^{2}+4\theta+2}{\left(\theta+2\right)^{2}},$$

$$\hat{h}_{\theta\theta}=\frac{4}{\left(\theta+2\right)^{3}},$$

$$\hat{p}_{\theta}=\frac{\beta}{\theta^{2}}-\frac{\alpha+1}{\theta},$$

$$\hat{L}_{\theta\theta}=-\frac{2n}{\theta^{2}}+\frac{n}{\left(1+\theta\right)^{2}},$$

$$\hat{\sigma}_{\theta\theta}=\frac{\theta^{2}\left(1+\theta\right)^{2}}{2n\left(1+\theta\right)^{2}-n\theta^{2}},$$

$$\hat{L}_{\theta\theta\theta} = \frac{4n}{\theta^3} - \frac{2n}{(1+\theta)^3},$$

Then,

$$\mathrm{P}^B_{\mathrm{Ent}} = \left[\begin{array}{c} \frac{\hat{\theta}(\hat{\theta}+1)}{\hat{\theta}+2} + \\ \left[\left[\frac{2}{(\hat{\theta}+2)^3} + \left(\frac{\hat{\theta}^2+4\hat{\theta}+2}{(\hat{\theta}+2)^2}\right)\left(\frac{\beta-\hat{\theta}(\alpha+1)}{\hat{\theta}^2}\right)\right] \frac{(\hat{\theta}(1+\hat{\theta}))^2}{2n(1+\hat{\theta})^2-n\hat{\theta}^2}\right] + \\ \left[\left[\frac{\hat{\theta}^2(1+\hat{\theta})^2}{2n(1+\hat{\theta})^2-n\hat{\theta}^2}\right]^2 \left[\frac{\hat{\theta}^2+4\hat{\theta}+2}{(\hat{\theta}+2)^2}\right]\left[\frac{2n}{\hat{\theta}^3} - \frac{n}{(1+\hat{\theta})^3}\right]\right]. \end{array} \right]^{-1}. \quad (3.23)$$

3.3 Simulation study

In this section, Monte Carlo simulation study is performed to compare the estimators by using mean square Errors (MSE's) as follows:

$$MSE\left(\hat{\mathrm{P}}^B_{\bullet}\right) = \frac{\sum_{i=1}^N \left(\hat{\mathrm{P}}^B_{\bullet} - \mu(\theta)\right)^2}{N}. \quad (3.24)$$

Where N is the number of replications. We generated 10000 samples of size n = 20, 40, 60 , 80 and 100 to represent small, moderate and large sample sizes from Lindley distribution with three values of θ $(\theta = 0.1,\ 1,\ 3)$.

In order to compare the Bayesian premium estimators obtained in the above section under three different loss functions and two priors, we choose the values of the extension of Jeffreys constants, $(c = 1, 2.5)$ and for the Inverted Gamma prior, the following pairs of values of the hyper parameters α and β are chosen

$(\alpha, \beta) = \{(1,\ 1.5), (1.5,\ 2)\}$, with two values of linex loss symmetry $(a = \pm 1)$ and $q = 1$ for entropy loss.

The results are summarized and tabulated in the following tables:

$\boldsymbol{\theta}$	**0.10**	**1.0**	**3.0**
$\boldsymbol{\mu}(\boldsymbol{\theta})$	**19.0909**	**1.5**	**0.416667**
n		**Ext.J.P**	
20	18.719562(0.137891)	1.483374(0.000276)	$0.414434(4.98628.10^{-6})$
40	18.705893(0.148230)	1.485517(0.000209)	$0.415701(9.33156.10^{-7})$
60	19.053733(0.001381)	$1.498388(2.59854.10^{-6})$	$0.415538(1.27464.10^{-6})$
80	18.938722(0.023158)	$1.493382(4.37979.10^{-5})$	$0.414043(6.88537.10^{-6})$
100	18.938133(0.023337)	$1.497925(4.30562.10^{-6})$	$0.416238(1.84041.10^{-7})$
n		**IG.P**	
20	17.804431(1.655002)	$1.495123(2.37851.10^{-5})$	$0.418247(2.49640.10^{-6})$
40	17.397672(2.867021)	$1.496145(1.48610.10^{-5})$	$0.41558(1.18156.10^{-6})$
60	17.491732(2.557338)	1.486513(0.000181)	$0.414668(3.99600.10^{-6})$
80	17.734848(1.838877)	$1.493021(4.87064.10^{-5})$	$0.415619(1.09830.10^{-6})$
100	18.597776(0.243171)	$1.499998(4.10^{-12})$	$0.413999(7.11822.10^{-6})$

Table 3.1 - Bayesian premium estimators and respective MSE's under squared error loss function $(\alpha = 1, \beta = 1.5, a = -1, c = 1)$.

$\boldsymbol{\theta}$	**0.10**	**1.0**	**3.0**
$\boldsymbol{\mu}\left(\boldsymbol{\theta}\right)$	**19.0909**	**1.5**	**0.416667**
n		**Ext.J.P**	
20	14.903964(17.53043)	$1.498348(2.72910.10^{-6})$	$0.415193(2.17267.10^{-6})$
40	17.834961(1.577383)	$1.498183(3.30148.10^{-6})$	$0.412039(2.14183.10^{-6})$
60	18.136684(0.910528)	$1.493726(3.93630.10^{-5})$	$0.416859(3.6864.10^{-8})$
80	18.607819(0.233367)	$1.492378(5.80948.10^{-6})$	$0.416303(1.3249.10^{-7})$
100	17.361210(2.991827)	1.487843(0.000147)	$0.415748(8.44561.10^{-7})$
n		**IG.P**	
20	15.850111(10.50271)	$1.496631(1.13501.10^{-5})$	$0.416006(4.36921.10^{-7})$
40	18.575453(0.265685)	1.487255(0.000162)	$0.416640(7.29.10^{-6})$
60	19.067037(0.000569)	1.489089(0.000119)	$0.413442(1.04006.10^{-7})$
80	18.548093(0.294639)	1.489417(0.000111)	$0.415829(7.02244.10^{-7})$
100	18.851246(0.057434)	$1.492716(5.30566.10^{-5})$	$0.416031(4.04496.10^{-7})$

Table 3.2 - Bayesian premium estimators and respective MSE's under squared error loss function ($\alpha = 1.5, \beta = 2, a = +1, c = 2.5$).

$\boldsymbol{\theta}$	**0.10**	**1.0**	**3.0**
$\boldsymbol{\mu}(\boldsymbol{\theta})$	**19.0909**	**1.5**	**0.416667**
n		**Ext.J.P**	
20	18.706146(0.148035)	1.484380(0.000243)	$0.415329(1.79024.10^{-6})$
40	18.908955(0.033103)	1.488878(0.000123)	$0.414208(6.04668.10^{-6})$
60	18.891763(0.039655)	$1.497278(7.40928.10^{-6})$	$0.415988(4.61041.10^{-6})$
80	18.908173(0.033389)	$1.498281(2.954961.10^{-6})$	$0.416496(2.9241.10^{-8})$
100	19.026982 (0.100322)	1.487455(0.000157)	$0.415179(2.21414.10^{-6})$
n		**IG.P**	
20	17.610532(2.191489)	1.486857(0.000172)	$0.414445(4.93728.10^{-6})$
40	17.574276(2.300148)	$1.493122(4.73068.10^{-5})$	$0.415956(5.05521.10^{-7})$
60	17.820992(1.612666)	$1.493434(4.31123.10^{-5})$	$0.415639(1.05678.10^{-6})$
80	17.945180(1.312674)	$1.492088(6.25997.10^{-5})$	$0.415879(6.20944.10^{-6})$
100	18.833766(0.066117)	$1.497639(5.57432.10^{-6})$	$0.416304(1.31769.10^{-7})$

Table 3.3 - Bayesian premium estimators and respective MSE's under linex loss function ($\alpha = 1, \beta = 1.5, a = -1, c = 1$).

$\boldsymbol{\theta}$	**0.10**	**1.0**	**3.0**
$\boldsymbol{\mu}(\boldsymbol{\theta})$	**19.0909**	**1.5**	**0.416667**
n		**Ext.J.P**	
20	15.975938(9.702988)	$1.495179(2.32420.10^{-5})$	$0.416206(2.12521.10^{-7})$
40	18.279793(0.657894)	$1.498343(2.74564.10^{-6})$	$0.414188(6.14544.10^{-6})$
60	18.067815(1.046703)	$1.496371(1.31696.10^{-5})$	$0.415223(2.08513.10^{-6})$
80	17.933340(1.339945)	$1.493254(4.55085.10^{-5})$	$0.416641(6.76000.10^{-10})$
100	18.249427(0.708076)	$1.492480(5.65504.10^{-5})$	$0.416360(9.4249.10^{-8})$
n		**IG.P**	
20	17.787725(1.698265)	1.489921(0.000101)	$0.413895(7.68398.10^{-6})$
40	18.685484(0.164362)	1.510593(0.000112)	$0.415405(1.59264.10^{-6})$
60	18.020992(1.144703)	$1.502698(7.27920.10^{-6})$	$0.415966(4.91401.10^{-7})$
80	17.400625(2.85703)	1.486900(0.000171)	$0.415153(2.29219.10^{-6})$
100	18.784024(0.094172)	$1.495416(2.10130.10^{-5})$	$0.415968(4.88601.10^{-7})$

Table 3.4 - Bayesian premium estimators and respective MSE's under linex loss function ($\alpha = 1.5, \beta = 2, a = +1, c = 2.5$).

$\boldsymbol{\theta}$	**0.10**	**1.0**	**3.0**
$\boldsymbol{\mu(\theta)}$	**19.0909**	**1.5**	**0.416667**
n		**Ext.J.P**	
20	17.343305(3.054088)	$1.503618(1.30899.10^{-6})$	$0.414691(3.90457.10^{-6})$
40	17.622991(2.154757)	$1.496511(1.21731.10^{-6})$	$0.4185030(3.37089.10^{-6})$
60	18.140113(0.903995)	1.486606(0.000179)	$0.414024(6.98544.10^{-6})$
80	17.530214(2.435741)	1.481438(0.000344)	$0.413987(7.1824.10^{-6})$
100	18.431041(0.435413)	$1.493254(4.55085.10^{-6})$	$0.416823(2.4336.10^{-6})$
n		**IG.P**	
20	18.361122(0.532575)	1.481980(0.000324)	$0.413554(9.69076.10^{-6})$
40	17.587366(2.260614)	1.473898(0.000681)	$0.415305(1.85504.10^{-6})$
60	17.728454(1.856259)	$1.495907(1.67526.10^{-5})$	$0.412041(2.13998.10^{-5})$
80	18.701742(0.151443)	$1.492903(5.03674.10^{-5})$	$0.416335(1.10224.10^{-7})$
100	17.815585(1.626428)	$1.490392(9.23136.10^{-5})$	$0.414688(3.91644.10^{-6})$

Table 3.5- Bayesian premium estimators and respective MSE's under entropy loss function ($\alpha = 1, \beta = 1.5, q = 1, c = 1$).

Remarks

This study deals with the Bayesian estimation problem based on Lindley distribution as a conditional distribution. Most authors used squared error as symmetric loss function. However, in practice, the real loss function is often not symmetric.

Simulation study revealed that the Bayesian premium estimator under entropy loss is efficient and realistic in most of the situation.

Furthermore, MSE of the Bayesian premium estimators for the entropy loss has a small values as compared with the corresponding Bayesian estimators under linex and squared error loss functions.

It may be noted here that when θ increases, the individual premium $\mu(\theta)$ decreases and the Bayesian premium estimator tends to $\mu(\theta)$.

Under the two above priors, we conclude that the performance is approximately equal with smaller posterior risk as compared.

From the above mentioned discussion, we may conclude that the Bayes procedure discussed in this paper can be recommended for their use.

Chapter 4

Conclusion and Perspectives

In this thesis, since the risk parameter for a policyholder is never known, we constructed Bayesian premium estimators under different loss functions. By imposing a prior distribution, we are able to describe the risk structure for the entire rating class. In practice, the choice of this prior distribution is subjective to personal judgements or induced from historical data of the corresponding group.

Using numerical simulation, it has been observed that the Bayesian premiums verified the condition of convergence to the individual premium which proves the efficacity of these estimators under different loss functions.

For future studies, we aim to verify the previous derivation using numerical algorithms like Metropolis-Hasting and Gibbs sampling.

In addition, we will search techniques to find linear credibility premiums under the entropy loss.

Bibliography

[1] Ahn, S.E., C.S., Park, H.M., Kim. Hazard rate estimation of a mixture model with censored lifetimes, Stoch. Environ. Res. Assess. 21 (2007), 711–716.

[2] Al-Kutubi, H.S. and Ibrahim, N. A. Bayes estimator for exponential distribution with extension of Jeffrey prior information, Malaysian Journal of mathematical Sciences, 3(2): 297-313, 2009.

[3] Bailey, A.L., Credibility procedures, Laplace's generalization of Bayes' rule and the combination of collateral knowledge with observed data, Proceedings of the Casualty Actuarial Society, vol. 37, 7–23, 1950.

[4] Basu, A. P., and Ebrahimi, N. Bayesian approach to life testing and reliability estimation using asymmetric loss function, J. Statist. Plann. Infer., 29, 21–31, 1991.

[5] Berger, J. O., Statistical decision theory and Bayesian analysis (2^{nd} edition), Springer-Verlag Inc (Berlin; New York), 1985.

[6] Boratynska, A., Posterior regret Γ-minimax estimation of insurance premium in collective risk model, Astin Bulletin, 38, (01), 277-291, 2008.

[7] Bühlmann, H., Experience rating and credibility, Astin Bulletin, vol. 4, p. 199–207, 1967.

[8] Bühlmann, H., Experience rating and credibility, Astin Bulletin, vol. 5, p. 157–165, 1969.

[9] A. Dermoune, P. Heinrich. A small step towards the hydrodynamic limit of a colored disordered lattice gas. C. R. Math. Acad. Sci. Paris **339**: no. 7, 507–511(2004).

[10] Bühlmann, H., Straub, E., Glaubgwürdigkeit für Schadensutze. Bulletin of the Swiss association of actuaries 70, 111–133, 1996.

[11] Calabria, R. and Pulcini, G., Point estimation under asymmetric loss functions for left truncated exponential samples, Comm. in Statist. – Theory and Methods, 25(3): 585- 600, 1969.

[12] De Vylder, F. Parameter estimation in credibility theory, Astin Bulletin, vol. 10, p. 99–112, 1978.

[13] De Vylder, F. and M. J. Goovaerts., Optimal parameter estimation under zero-excess assumptions in a classical model, Insurance : Mathematics and Economics, vol. 11, p. 1–6, 1992a.

[14] Dey, D., Ghosh, M., Strawderman, W.E. On estimation with balanced loss functions, Statistics & Probability Letters, 45, 97-101, 1999.

[15] Dey, D.K., and Liu, P.L. On comparison of estimators in a generalized life model, Microelectron Reliab., 32, 207 – 221, 1992.

[16] Diaconis. P, Ylvisaker. D. Conjugate priors for exponential families, The Annals of Statistics, Vol. 7, No. 2, 269-281, 1979.

[17] Dubey, A. et A. Gisler., On parameter estimation in credibility, Bulletin of the Swiss association of actuaries, vol. 81, p. 187–211, 1981.

[18] Feller, W. An introduction to probability theory and its applications, Volume II, Wiley, New York, p.258, 1966.

[19] Garthwaite, H. P., Kadane, B. J. and O'Hagan, A. Elicitation, working paper in university of Sheffield, 2004.

[20] Gauss, C.F. Least squares method for the combinations of observations, (translated by J. Bertrand 1955). Mallet-Bachelier, Paris, 1810.

[21] Gerber, H. U. An introduction to mathematical risk theory, Huebner Foundation, Philadelphia, 1979.

[22] Ghitany, M.E, Al-Mutairi D.K. Estimation methods for the discrete Poisson–Lindley distribution. Journal of Statistical Computation and Simulation,79(1):1–9, 2009.

[23] Ghitany, M.E, Atieh, B., Nadarajah, S. Lindley distribution and its application. Mathematics and computers in simulation, 78:493–506, 2008.

[24] Goel, P.K., On implications of credible means being exact Bayesian, Scandinavian Actuarial Journal, p. 41–46, 1982.

[25] Goovaerts, M. J. and W. J. Hoogstad., Credibility Theory,4, Surveys of actuarial studies, Nationale-Nederlanden N.V., Netherlands, 1987.

[26] Goulet, V., Théorie de la crédibilité : histoire, principes et applications, mémoire de maîtrise, Université Laval, 1994.

[27] Hachemeister, C.A. Credibility for regression models with application to trend, Credibility theory and applications, Proceedings of the Berkeley actuarial research conference on credibility, Academic Press, New York, p. 129–163, 1975.

[28] Howlader, H.A., Hossain, A. Bayesian survival estimation of Pareto distribution of the second kind based on failure censored data, Computational Statistics Data Analysis, 38:301–314, 2002.

[29] Herzog, T.N. Introduction to credibility theory, 2nd ed. Actex Publications, Winsted, 1996.

[30] N.A, Ibrahim, C.B, Guure, A.O.M, Ahmed, Bayesian estimation of two-parameter weibull distribution using extension of Jeffreys' prior information with three loss functions, Mathematical Problems in Engineering, 2012.

[31] Jaheen, Z.F. On record statistics from a mixture of two exponential distributions, J. Statist. Comput. Simul., 75, 1-11, 2005.

[32] Jewell, W.S. Credible means are exact bayesian for exponential families, Astin Bulletin, vol. 8, p. 77–90, 1974.

[33] Kaas, R., Dannenburg, D., Goovaerts, M. Exact credibility for weighted observations, Astin Bulletin 27 (2), 287–295, 1996.

[34] Klugman, S.A., H. Panjer, and G. Willmot. Loss Models: From data to decisions, 2^{nd} edition, Wiley, New York, Chapter 16, 2004.

[35] Klugman, S.A., H. Panjer, and G. Willmot, Loss Models: From data to decisions, 3^{rd} ed, Wiley, New York, 2008.

[36] Kiapour, A., and Nematollahi. Robust Baysian prediction and estimation under squared log error loss function, Statist. Probab. Lett., 81: 1717-1724, 2011.

[37] H. Krishna, K. Kumar. Reliability estimation in Lindley distribution with progressively type II right censored sample, Math. Comput. Simul, 82 (2), 281–294, 2011.

[38] Legendre, A. Nouvelles méthodes pour la determination des orbites des comètes, Courcier, City Paris, 1805.

[39] D.V., Lindley. Fiducial distributions and Bayes' theorem, Journal of the Royal Statistical Society, Series A20,102–107, 1958.

[40] Lindley, DV. Approximate Bayes methods. Trabajos de Estadistica de Investigacion Operativa, 31(1):223–245, 1980

[41] Makov, U., and Landsman, Z. Exponential dispersion models and credibility, Scandinavian actuarial journal, 1, 89-96, 1998a.

[42] Mayerson, A.L. A Bayesian view of credibility, Proceedings of the Casualty Actuarial Society, vol. 51, 85–104, 1964.

[43] Mowbray, A.H. How extensive a payroll exposure is necessary to give a dependable pure premium?, PCAS I, 24-30, 1914.

[44] Nassar, M., and Eissa, F.H. Bayesian estimation for the exponentiated Weibull model, Commun Statistical theory Meth., 33(10), 2343-2362, 2004.

[45] Norberg, R. The credibility approach to ratemaking, Scandinavian actuarial journal, 181–221, 1979.

[46] R. Norberg. Credibility theory, encyclopedia of actuarial science, Wiley, Chichester, UK, 2004.

[47] Pandey, B.N. T-estimator of the scale parameter of the exponential distribution using Linex loss function. Commun. Statistical Theory. Meth., 26(9), 2191–2202, 1997.

[48] Parsian, A., and Kirmani. Estimation under Linex loss function. Handbook of ap-

plied econometrics and statistical inference, Statistics Textbooks Mono- graphs, 53-76, 2002.

[49] Rojo. On the admissibility of "cx+d" with respect to the Linex loss Function, Commun. Statistical Theory. Meth, 116, 3745-3748, 1987.

[50] Sajid Ali, Muhammad Aslam, Syed Mohsin Ali Kazmi. A study of the effect of the loss function on Bayes Estimate, posterior risk and hazard function for Lindley distribution, Applied Mathematical Modelling, 37, 6068–6078, 2013.

[51] Sankaran, M. The discrete Poisson–Lindley distribution. Biometrics, 26, 145-149, 1970.

[52] Soliman, A. Comparison of Linex and quadratic Bayes estimators for the Rayleigh distribution. Commun. Statist. Theor. Meth., 29(1), 95–107, 2000.

[53] Thompson, R.D. and Basu, A.P. Asymmetric loss function for estimating reliability. Journal of Bayesian Analysis in Statistics and Econometrics, Wiley, New York, 1996.

[54] Varian, H.R. A Bayesian approach to real estate assessment, Studies in Bayesian Econometrics and Statistics, Amsterdam: North-Holland, 195–208, 1975.

[55] Whitney, A.W. The theory of experience rating, Proceedings of the Casualty Actuarial Society, vol. 4, 275–293, 1918.

[56] Zamani, H., Ismail, N. Negative Binomial–Lindley distribution and its application. Journal of mathematics and Statistics, 6(1), 4–9, 2010.

[57] Zeghdoudi, H, Nedjar, Gamma Lindley distribution and its application. Journal of Applied probability and statistics, vol. 10, 2015.

[58] Zellner, A. Bayesian estimation and prediction using asymmetric loss functions. Journal of Amsterdam statistical association, 81, 446–451, 1986.

Annexe

HACETTEPE UNIVERSITY
Journal of Mathematics and Statistics

EARLY ACCESS

Statistics

On weighted balanced loss function under the Esscher principle and credibility premiums
Farouk Metiri, Halim Zeghdoudi, Mohamed Riad Remita
Doi: 10.15672/HJMS.201612216466
FULL TEXT

Generalized Multi-Phase Regression-type Estimators Under the Effect of Measuemnent Error to Estimate the Population Variance
Saadia Masood, Javid Shabbir
Doi: 10.15672/HJMS.201612116352
FULL TEXT

Different estimation methods and joint condence regions for the parameters of a generalized inverted family of distributions
Wenhao Gui, Lei Guo
Doi: 10.15672/HJMS.201612016315
FULL TEXT

On weighted balanced loss function under the Esscher principle and credibility premiums

Farouk Metiri, Halim Zeghdoudi and Mohamed Riad Remita

Abstract

This paper focuses on weighted balanced loss function under the Esscher principle (WBLF) of which we explore the modern practice of credibility theory and we generalize credibility premiums by using the WBLF.We obtain a distribution-free approach under the WBLF and the Esscher premium by using a minimization technique. Also, we discuss the consistency of the credibility premium generated by this distribution-free approach.

Keywords: Credibility premium, Esscher premium, loss function.

2000 AMS Classification: 62P05, 62.1

1. Introduction and motivation

According to Rodermund (1989), the concept of credibility has been the casualty actuaries most important and enduring contribution to casualty actuarial science.

In this sense, credibility theory is used to determine the expected claims experience of an individual risk when those risks are not homogeneous, given that the individual risk belongs to a heterogeneous collective. The main objective of the credibility theory is to calculate the weight which should be assigned to the individual risk data to determine a fair premium to be charged, for recent detailed introductions to credibility theory, see Norberg (2004), Bühlmann and Gisler (2005).

Moreover, the credibility assumed that the individual risk, X, has a density $f(x \mid \theta)$ indexed by a parameter $\theta \in \Theta$ which has a prior distribution with density $\pi(\theta)$. Let, now, $\pi^x(\theta)$ be the posterior density when x is observed. In actuarial science, the unknown risk premium $P_R^L \equiv P_R^L(\theta)$ is obtained by minimizing the expected loss $E_f[L(\theta, P)]$, with $p \in P$ for some loss function L. If experience is not available, the actuary chooses the collective premium P_C^L, which is given by minimizing the risk function, i.e. $E_\pi[L(P_R^L(\theta), P_C^L)]$.
In addition, if experience is available, the actuary takes a sample X from the random variables X_i, $i = \overline{1,t}$ and uses this information to the unknown risk premium $P_R^L(\theta)$, through the Bayes premium P_B^L, obtained by minimizing the Bayes risk, i.e. $E_{\pi^x}[L(P_R^L(\theta), P_B^L)]$.

According to Heilmann (1989), many credibility premiums were obtained under statistical decision theory from a Bayesian point of view and using the weighted squared error loss function (WLF henceforth), $L_1(P,x) = h(x)(x-P)^2$, using different functional forms of $h(x)$ we have different premium principles (such as net premium principle, expected value premium principle, variance premium principle, standard deviation premium principle, proportional hazards premium principle, principle of equivalent utility, dutch premium principle, Wang's premium principle, exponential principle, mean value principle, zero utility principle, Swiss premium calculation principle, Orlicz principle, Esscher principle). For example, if we take $h(x) = 1$ and $h(x) = e^{hx}$, $h > 0$, we have the net and the Esscher premium principles, see Heilmann (1989), Gómez (2006), and others.

In today's point of view, it would be better to understand credibility premium as a simplified version of Bayes estimation of the individual pure premium. It is well known that the credibility premium can be written as a convex combination between the individual and the collective information. Under the case of the exponential family of distributions, exactly in the case of the pair: Poisson- Gamma, the Bayes Esscher premium can be written as a credibility formula in the form:

$$P_B^{L_1} = Z\left(t\right) g\left(\bar{x}\right) + \left(1 - Z\left(t\right)\right) P_C^{L_1}$$

with (see Heilmann (1989) and Gómez (2006) for details):

$P_B^{L_1}$: the Bayes premium obtained under WLF;
$P_C^{L_1}$: the collective premium obtained under WLF;
$g\left(\bar{x}\right)$: a function of the observed data;
$Z\left(t\right)$: is the credibility factor, satisfying the condition $0 \leq Z\left(t\right) \leq 1$.

In this work, we use the weighted balanced loss function (WBLF) to obtain new credibility premiums, WBLF is a generalized loss function introduced by Zellner (1994) (see Gupta and Berger (1994), pp.371-390) and which appears also in Dey et al. (1999) and Farsipour and Asgharzadhe (2004). It is given by

$$L_2\left(P, x\right) = \omega h(x)(\delta_0(x) - P)^2 + \left(1 - \omega\right) h(x)(x - P)^2$$

where $0 \leq \omega \leq 1$, $h(x)$ is a positive weight function, and $\delta_0(x)$ is a function of the observed data (see Jafari et al. (2006)). When ω is chosen to equal 0,This loss includes as a particular case the WLF, i.e.

$$\begin{aligned} L_2\left(P, x\right) &= 0h(x)(\delta_0(x) - P)^2 + \left(1 - 0\right) h(x)(x - P)^2 \\ &= h(x)\left(x - P\right)^2 = L_1(P, x) \end{aligned}$$

Moreover, our work is a generalization of Gómez Déniz (2008) and the results obtained here are very close to those obtained by Najafabadi et al. (2010) whose approximate the Bayes estimator with respect to a general loss function and general prior distribution by a convex combination of the observation mean and mean of prior, say, approximate credibility formula.

The paper is organized as follows. Section 2 describes the Esscher premium principle and its properties . Section 3 is dedicated to derive the Esscher credibility premiums under WBLF. Section 3 provides the main contribution of this work, i.e., the solutions under the distribution free approach. Finally, a small simulation is carried out to illustrate the theoretical conclusions and some remarks.

2. Properties of the Esscher premium principle

Let χ denote the set of non-negative random variables on the probability space (Ω, F, P). Goovaerts et al. (1984) describe the Esscher premium as the expected value of the risk X after multiplying the density of X by an increasing weight function, which of course makes the risk less attractive to the insurer. The Esscher premium of $X \in \chi$ is given by

$$H\left[X\right] = e^{hx}\left(x - P\right)^2, \ h > 0.$$

which h reflects the risk averseness of the insurer. In fact, the distribution function F_x of X is replaced by its Esscher transform, denoted by $F_{X,h}$, where h is a real parameter:

$$dF_{X,h}\left(x\right) = \frac{e^{hx} dF_X\left(x\right)}{\int_0^\infty e^{hx} dF_X\left(x\right)}$$

Clearly, $F_{X,h}$ is also a distribution function. So the Esscher premium of X, with parameter h can be calculated as

$$H[X] = \int_0^\infty x dF_{X,h}\left(x\right)$$

Some properties of Esscher premium principle are listed as follows. The proofs can be easily checked.

- **Risk loading**: $H[X] > E\left[X\right]$ for all $X \in \chi$, and $h > 0$. In addition, when $h \to 0$, we have $H[X] \to E(X)$ which is equal to the net premium principle. Loading for risk is desirable because one generally requires a premium rule to charge at least the expected payout of the risk X, namely $E(X)$, in exchange for insuring the risk. Otherwise, the insurer will lose money on average.
- **No unjustified risk loading**: If a risk $X \in \chi$ is identically equal to a constant $c \geq 0$ (almost everywhere), then $H[c] = e^{hc}\left(c - P\right)^2 = c$. If we know for certain (with probability 1) that the insurance payout is c, then we have no reason to charge a risk loading because there is no uncertainty as to the payout.
- **Maximal loss (or no rip-off)**: $H[X] \leq esssup[X]$ for all $X \in \chi$.
- **Translation equivariance (or translation invariance)**: $H[X + c] = H[X] + c$ for all $X \in \chi$ and all $c \geq 0$. If we increase a risk X by a fixed amount c, then the premium for $X + c$ should be the premium for X increased by the fixed amount c.
- **Additivity for independent risks**: If $X, Y \in \chi$ are independent of each other, then $H[X + Y] = H[X] + H[Y]$.
- **Monotonicity**: If $X(\omega) \leq Y(\omega)$ for all $\omega \in \Omega$, then $H[X] \leq H[Y]$.
- **Preserving of first stochastic dominance (FSD) ordering**: If $S_X(t) \leq S_Y(t)$ for all $t \geq 0$, then $H[X] \leq H[Y]$.
- **Preserving of stop-loss (SL) ordering**: If $E[X - d]_+ \leq E[Y - d]_+$ for all $d \geq 0$, then $H[X] \leq H[Y]$.
- **Continuity**: Let $X \in \chi$, then $lim_{a \to 0^+}\ H[max(X - a, 0)] = H[X]$, and $lim_{a \to \infty} H[min(X, a)] = H[X]$.

Now, we give the individual premium, the collective premium and the Bayesian premium under the Esscher principle:

The individual premium
The individual premium of risk X and parameter θ under Esscher premium principle is

$$P_R^L = E_f\left[L(\theta, P)\right] = \frac{E_f\left[Xe^{hx} \mid \theta\right]}{E_f\left[e^{hx} \mid \theta\right]}$$

The collective premium
If we have no information, the actuary charges the collective premium to the insured which is given by

$$P_C^L = \frac{E_\pi\left[P_R^L e^{hP_R^L}\right]}{E_\pi\left[e^{hP_R^L}\right]}$$

The Bayesian premium
To calculate the Bayesian premium, we use both the prior information about the parameters of the loss process and the actual loss experience observed during the policy period. The posterior density function is obtained using the prior density function and the data on actual losses from the Bayes' theorem. Let $f(x \mid \theta)$ be the probability function of X, and let $\pi(\theta)$ denote the prior density function of θ. Let $\pi^X(\theta)$ be the posterior density function, then:

$$P_B^L = E_{\pi^x}\left[L(P_R^L, P_B^L)\right] = \frac{E_{\pi^x}\left[P_R^L e^{hP_R^L}\right]}{E_{\pi^x}\left[e^{hP_R^L}\right]}$$

3. Derivation of premiums under WBLF

In this section, we aim to use the WBLF to derive a new credibility formula under the Esscher premium principle. Next lemma is a generalization of Lemma 3.1 in Jafari et al. (2006).

3.1. Lemma. *Under WBLF and prior π, the risk , collective and Bayes premium are given by*

$$P_R^{L_2} = \omega \frac{E_{f(x|\theta)}\left[\delta_0(x)h(x)\,|\theta\right]}{E_{f(x|\theta)}\left[h(x)\,|\theta\right]} + (1-\omega)\frac{E_{f(x|\theta)}\left[Xh(x)\,|\theta\right]}{E_{f(x|\theta)}\left[h(x)\,|\theta\right]}$$

$$\begin{aligned} P_C^{L_2} &= \omega \frac{E_\pi\left[\delta_0(x)h(P_R^{L_2})\right]}{E_\pi\left[h(P_R^{L_2})\right]} + (1-\omega)\frac{E_\pi\left[P_R^{L_2}h(P_R^{L_2})\right]}{E_\pi\left[h(P_R^{L_2})\right]} \\ &= \omega\delta_0^* + (1-\omega)\frac{E_\pi\left[P_R^{L_2}h(P_R^{L_2})\right]}{E_\pi\left[h(P_R^{L_2})\right]} \end{aligned}$$

$$P_B^{L_2} = \omega\delta_0^* + (1-\omega)\frac{E_{\pi^x}\left[P_R^{L_2}h(P_R^{L_2})\right]}{E_{\pi^x}\left[h(P_R^{L_2})\right]},$$

where δ_0^* *is a target estimator for the risk (individual) premium* $P_R^{L_2}$*.*

Proof. The proof, which is similar to the one given by Dey et al. (1999).
Under WBLF, we minimize $E_{f(x|\theta)}[L_2(\theta, P_R^{L2})]$ with respect to P_R^{L2}.
Under WBLF , we minimize $E_{\pi(\theta)}[L_2(P_R^{L_1}, P_C^{L_2})]$ with respect to $P_C^{L_2}$.
We replace $\pi(\theta)$ by $\pi^x(\theta)$ to obtain the Bayes premium $P_B^{L_2}$. □

3.2. Lemma. *If the Bayes premium obtained under $L_1(P,x)$ is a credibility formula, the Bayes balanced premium obtained under WBLF is also a credibility formula in this form:*

$$P_B^{L_2} = \omega\delta_0^* + (1-\omega)\frac{E_{\pi^x}\left[\frac{e^{\frac{h}{h-\theta}}}{h-\theta}\right]}{E_{\pi^x}\left[e^{\frac{h}{h-\theta}}\right]}$$

Proof. Consider the case in which, the claim follows a Poisson distribution with parameter $\theta > 0$ and the prior is a gamma distribution $\pi(\theta) \propto \theta^{\alpha-1}e^{-\beta\theta}$, $\alpha > 0, \beta > 0$. Suppose also that the actuary chooses the WLF to obtain the Esscher risk premium and the $WBLF$ to obtain the Esscher collective and Bayes premiums. Then, we have:

$$P_R^{L_1} = \frac{E_{f(x|\theta)}\left[Xe^{hx}/\theta\right]}{E_{f(x|\theta)}\left[e^{hx}/\theta\right]} = \theta e^h$$

$$P_C^{L_2} = \omega\delta_0^* + (1-\omega)\frac{\alpha e^h}{\beta - he^h}$$

$$P_B^{L_2} = \omega\delta_0^* + (1-\omega)\frac{(\alpha+t)e^h}{\beta+t-he^h} = z(t)\,l\left(P_C^{L_2}\right) + (1-z(t))\,l(e^h\bar{x})$$

where $z(t) = \dfrac{\beta - h^h}{\beta+t-he^h}$ and $l(x) = \omega\delta_0^* + (1-\omega)x$.
However, if we replace Poisson(θ) with the exponential distribution Exp(θ), $\theta > 0$, $P_B^{L_2}$ no longer has a credibility formula, because we have

$$P_R^{L_1} = \frac{1}{h-\theta},$$

$$P_C^{L_2} = \omega\delta_0^* + (1-\omega)\frac{E_{\pi}\left[\frac{e^{\frac{h}{h-\theta}}}{h-\theta}\right]}{E_{\pi}\left[e^{\frac{h}{h-\theta}}\right]},$$

and

$$P_B^{L_2} = \omega\delta_0^* + (1-\omega)\frac{E_{\pi^x}\left[\frac{e^{\frac{h}{h-\theta}}}{h-\theta}\right]}{E_{\pi^x}\left[e^{\frac{h}{h-\theta}}\right]},$$

which is not a credibility formula. □

3.3. Remark. *Under the exponential family, the Bayes balanced premium is linear only under the case Poisson-gamma.*

4. The distribution-free approach: Main results

According to Bühlmann (1967), the classical formula in credibility theory often calculates the premium as a weighted sum of the average experience of the policyholder and the average experience of the entire collection of policyholders. The main idea in this work is to change the exact credibility premium $H(\mu(\theta)|X_1, X_2, ..., X_n)$ by a linear expression of the form $c_0 + c_1 H_n[x]$ which $H_n[x] = \frac{\sum_{i=1}^n X_i e^{hX_i}}{\sum_{i=1}^n e^{hX_i}}$ indicates the empirical Esscher premium, depending on the past claims $X_i, i = \overline{1, n}$. Using the $WBLF$, we will suppose that the variables $X_1|\theta, X_2|\theta, ..., X_n|\theta$ are independently and identically distributed. To simplify the presentation, we use the following notations

$$\mu(\theta) = \omega \frac{E_{f(x|\theta)}\left[\delta_0(x)e^{hx}|\theta\right]}{E_{f(x|\theta)}\left[e^{hx}|\theta\right]} + (1-\omega)\frac{E_{f(x|\theta)}\left[Xe^{hx}|\theta\right]}{E_{f(x|\theta)}\left[e^{hx}|\theta\right]},$$

is the individual Esscher premium.

$$m = \omega\delta_0^* + (1-\omega)\frac{E_\pi\left[\mu(\theta)e^{h\mu(\theta)}\right]}{E_\pi\left[e^{h\mu(\theta)}\right]},$$

is the collective Esscher premium. Then, the coefficients c_0, c_1 must be determined by minimizing

(1)

$$\min_{c_0,c_1} E\left[\omega\left(\delta_0 - c_0 - c_1 H_n[x]\right)^2 e^{h\mu(\theta)} + (1-\omega)\left(\mu(\theta) - c_0 - c_1 H_n[x]\right)^2 e^{h\mu(\theta)}\right].$$

In order to find the solution to (1), we write

$$m_h(\theta) = E\left[e^{h\mu(\theta)} \mid \theta\right]$$

$$m_h = E\left[m_h(\theta)\right] = E_\Pi\left[e^{h\mu(\theta)}\right]$$

$$f_n(\theta) = E\left[H_n[x] \mid \theta\right]$$

$$E^*\left[f_n(\theta)\right] = \frac{E\left[f_n(\theta)\, m_h(\theta)\right]}{E\left[m_h(\theta)\right]} = \frac{E\left[H_n[x]\, e^{h\mu(\theta)}\right]}{E\left[e^{h\mu(\theta)}\right]},$$

where $\pi^*(\theta) = \dfrac{\pi(\theta)m_h(\theta)}{m_h}$ is a probability distribution function. To achieve the minimum in (6), the derivative with respect to "c_0" must be set to zero, namely,

$$E\left[\omega e^{h\mu(\theta)}\left(\delta_0 - c_0 - c_1 H_n[x]\right)\right] + E\left[(1-\omega)\, e^{h\mu(\theta)}\left(\mu(\theta) - c_0 - c_1 H_n[x]\right)\right] = 0$$

$$\omega\left(E\left[e^{h\mu(\theta)}\delta_0\right] - c_0 E\left[e^{h\mu(\theta)}\right] - c_1 E\left[e^{h\mu(\theta)} H_n[x]\right]\right) +$$

$$(1-\omega)\left(E\left[e^{h\mu(\theta)}\mu(\theta)\right]-c_0E\left[e^{h\mu(\theta)}\right]-c_1E\left[e^{h\mu(\theta)}H_n\left[x\right]\right]\right)=0$$

$$c_0E\left[e^{h\mu(\theta)}\right]=\omega E\left[e^{h\mu(\theta)}\delta_0\right]+(1-\omega)\,E\left[e^{h\mu(\theta)}\mu(\theta)\right]-c_1E\left[e^{h\mu(\theta)}H_n\left[x\right]\right]$$

$$\begin{aligned}
c_0 &= \frac{\omega E\left[e^{h\mu(\theta)}\delta_0\right]+(1-\omega)\,E\left[e^{h\mu(\theta)}\mu(\theta)\right]-c_1E\left[e^{h\mu(\theta)}H_n\left[x\right]\right]}{E\left[e^{h\mu(\theta)}\right]}\\
c_0 &= \omega\frac{E\left[\delta_0e^{h\mu(\theta)}\right]}{E\left[e^{h\mu(\theta)}\right]}+(1-\omega)\,\frac{E\left[\mu(\theta)e^{h\mu(\theta)}\right]}{E\left[e^{h\mu(\theta)}\right]}-c_1\frac{E\left[H_n\left[x\right]e^{h\mu(\theta)}\right]}{E\left[e^{h\mu(\theta)}\right]}\\
c_0 &= m-c_1\frac{E\left[H_n\left[x\right]e^{h\mu(\theta)}\right]}{E\left[e^{h\mu(\theta)}\right]}\\
c_0 &= m-c_1\frac{E\left[f_n\left(\theta\right)m_h\left(\theta\right)\right]}{E\left[m_h\left(\theta\right)\right]}
\end{aligned}$$

$$c_0=m-c_1E^*\left[f_n\left(\theta\right)\right]. \tag{8}$$

Now, the problem is equivalent to:

$$\min_{c_1}E\left[\begin{array}{c}\omega\left(\delta_0-m+c_1E^*\left[f_n\left(\theta\right)\right]-c_1H_n\left[x\right]\right)^2e^{h\mu(\theta)}\\+\left(1-\omega\right)\left(\mu(\theta)-m+c_1E^*\left[f_n\left(\theta\right)\right]-c_1H_n\left[x\right]\right)^2e^{h\mu(\theta)}\end{array}\right]=0$$

$$\min_{c_1}E\left[\begin{array}{c}\omega e^{h\mu(\theta)}\left(\delta_0-m-c_1\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)\right)^2\\+(1-\omega)e^{h\mu(\theta)}(\mu(\theta)-m-c_1\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right))^2\end{array}\right]=0$$

Taking derivative to "c_1":

$$\begin{aligned}
&2c_1E\left[\omega e^{h\mu(\theta)}\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2\right]-2E\left[\omega e^{h\mu(\theta)}\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)\left(\delta_0-m\right)\right]\\
&+2c_1E\left[\left(1-\omega\right)e^{h\mu(\theta)}\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2\right]\\
&-2E\left[\left(1-\omega\right)e^{h\mu(\theta)}\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)\left(\mu(\theta)-m\right)\right]=0\\
&c_1E\left[e^{h\mu(\theta)}\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2\right]=E\left[\omega e^{h\mu(\theta)}\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)\left(\delta_0-m\right)\right]\\
&\qquad+E\left[\left(1-\omega\right)e^{h\mu(\theta)}\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)\left(\mu(\theta)-m\right)\right]
\end{aligned}$$

$$
\begin{aligned}
c_1 &= \frac{E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)\left(\omega\delta_0-\omega\mu\left(\theta\right)+\mu\left(\theta\right)-m\right)e^{h\mu(\theta)}\right]}{E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2 e^{h\mu(\theta)}\right]}\\
c_1 &= \frac{E\left[\left(\omega\delta_0+(1-\omega)\mu\left(\theta\right)-m\right)\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)e^{h\mu(\theta)}\right]}{E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2 e^{h\mu(\theta)}\right]}\\
c_1 &= \frac{E\left[E\left[\left(\omega\delta_0+(1-\omega)\mu\left(\theta\right)-m\right)\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)e^{h\mu(\theta)}\mid\theta\right]\right]}{E\left[E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2 e^{h\mu(\theta)}\mid\theta\right]\right]}\\
c_1 &= \frac{E\left[E\left[\left(\omega\delta_0+(1-\omega)\mu\left(\theta\right)-m\right)\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)e^{h\mu(\theta)}\mid\theta\right]\right]}{E\left[E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2 e^{h\mu(\theta)}\mid\theta\right]\right]}\\
c_1 &= \frac{\frac{1}{m_h}E\left[E\left[\left(\omega\delta_0+(1-\omega)\mu\left(\theta\right)-m\right)\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)e^{h\mu(\theta)}\mid\theta\right]\right]}{\frac{1}{m_h}E\left[E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2 e^{h\mu(\theta)}\mid\theta\right]\right]}\\
c_1 &= \frac{\frac{1}{m_h}E\left[\left(\omega\delta_0+(1-\omega)\mu\left(\theta\right)-m\right)\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)m_h\left(\theta\right)\right]}{\frac{1}{m_h}E\left[E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2 m_h\left(\theta\right)\right]\right]}\\
c_1 &= \frac{E\left[\omega(\delta_0-\delta_0^*)+(1-\omega)(\mu\left(\theta\right)-\frac{E_\pi\left[\mu(\theta)e^{h\mu(\theta)}\right]}{E_\pi\left[e^{h\mu(\theta)}\right]})\left(f_n\left(\theta\right)-E^*\left[f_n\left(\theta\right)\right]\right)\frac{m_h(\theta)}{m_h}\right]}{E\left[E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2\frac{m_h(\theta)}{m_h}\right]\right]}\\
c_1 &= \frac{(1-\omega)E\left[(\mu\left(\theta\right)-\frac{E_\pi\left[\mu(\theta)e^{h\mu(\theta)}\right]}{E_\pi\left[e^{h\mu(\theta)}\right]})\left(f_n\left(\theta\right)-E^*\left[f_n\left(\theta\right)\right]\right)\frac{m_h(\theta)}{m_h}\right]}{E\left[E\left[\left(H_n\left[x\right]-E^*\left[f_n\left(\theta\right)\right]\right)^2\frac{m_h(\theta)}{m_h}\right]\right]}\\
c_1 &= \frac{(1-\omega)E\left[(\mu\left(\theta\right)-\frac{E_\pi\left[\mu(\theta)e^{h\mu(\theta)}\right]}{E_\pi\left[e^{h\mu(\theta)}\right]})\left(f_n\left(\theta\right)-E^*\left[f_n\left(\theta\right)\right]\right)\frac{m_h(\theta)}{m_h}\right]}{var^*\left[f_n\left(\theta\right)\right]+E^*\left[var\left[H_n\left[x\right]\mid\theta\right]\right]}
\end{aligned}
$$

(9) $$c_1=\frac{(1-\omega)cov^*\left(\mu(\theta),f_n\left(\theta\right)\right)}{var^*\left[f_n\left(\theta\right)\right]+E^*\left[var\left[H_n\left[x\right]\mid\theta\right]\right]}.$$

Thus,

$$
\begin{aligned}
H(\mu(\theta)|X_1,X_2,...,X_n) &= c_0+c_1H_n\left[x\right]\\
&= m-c_1E^*\left[f_n\left(\theta\right)\right]+c_1H_n\left[x\right]\\
&= c_1H_n\left[x\right]+\left(1-\frac{c_1E^*\left[f_n\left(\theta\right)\right]}{m}\right)m
\end{aligned}
$$

5. Numerical simulation

This section is made in order to illustrate the convergence of the empirical premium to the individual Esscher premium using a numerical simulation. We assume that X follows a Poisson distribution with parameter θ, and the prior is a gamma distribution. Taking $\delta_0(x)=\bar{x}e^h$

$$H(\mu(\theta)|X_1, X_2, ..., X_n) = c_1 H_n[x] + \left(1 - \frac{c_1 E^*[f_n(\theta)]}{m}\right) m$$

With:

$$c_1 = \frac{(1-\omega) cov^*(\mu(\theta), f_n(\theta))}{var^*[f_n(\theta)] + E^*[var[H_n[x] \mid \theta]]}$$

It is quite difficult to work out a closed form of c_1 due to the obstacle in the analytic calculation of $f_n(\theta) = E[H_n[x] \mid \theta]$ where $H_n[x] = \frac{\sum_{i=1}^{n} X_i e^{hX_i}}{\sum_{i=1}^{n} e^{hX_i}}$. Thus,

instead, we use a Monte Carlo method to compute numerically c_1. The algorithm is described as follows:

1- Randomly sample 4 values, θ_k, $k = 1, 2, 3, 4$ from distribution with density $\pi^*(\theta) \sim gamma(\alpha, \beta - he^h)$.

2- For each θ_k, we produce 1000 repetitions of sampling data, each of which consists of n independent and identically distributed values.

3- For each θ_k, we find the vector H_j (the empirical premium), according to this vector, we calculate: U_k, V_k, W_k ,i.e., compute:

$$H_j = \frac{\sum_{i=1}^{n} X_{ij} e^{hX_{ij}}}{\sum_{i=1}^{n} e^{hX_{ij}}}$$

$$U_k = \frac{\sum_{j=1}^{1000} H_j}{1000}$$

$$V_k = \frac{\sum_{j=1}^{1000} (H_j - U_k)^2}{1000 - 1}$$

$$W_k = \frac{\sum_{j=1}^{1000} \sum_{i=1}^{n} X_{ij} e^{hX_{ij}}}{\sum_{j=1}^{1000} \sum_{i=1}^{n} e^{hX_{ij}}}.$$

4- We calculate:

$$\begin{aligned}
A &= \frac{1}{4-1} \sum_{k=1}^{4} \left(U_k - \bar{U}\right)\left(W_k - \bar{W}\right) = cov^*(\mu(\theta), f_n(\theta)) \\
B &= \frac{1}{4-1} \sum_{k=1}^{4} \left(U_k - \bar{U}\right)^2 = var^*[f_n(\theta)] \\
C &= \frac{1}{4} \sum_{k=1}^{4} V_k = E^*[var[H_n[x] \mid \theta]] \\
D &= \frac{1}{4} \sum_{k=1}^{4} U_k = E^*[f_n(\theta)]
\end{aligned}$$

then,

$$c_1 = \frac{(1-\omega)\,A}{B+C}.$$

Simulation I

we take $h = 0.8$, $\omega = 0.9$, $\alpha = 2$ and $\beta = 6$. In addition, four different values of θ are given. Furthermore, two sample sizes are considered: $n = 100$ and $n = 150$. The corresponding simulation results are listed in the following tables:

n=100	$\mu(\theta)$	c_1	m=$P_c^{L_2}$	$\bar{H}$	$sd_{H(\mu(\theta)\mid X_1,X_2,\ldots,X_n)}$
θ=0.2	0.445108	0.0880289	0.503881	0.4467347	0.01009818
θ=0.4	0.890216	0.0880289	0.903917	0.8855569	0.00539665
θ=0.6	1.335300	0.0880289	1.312647	1.332238	0.2863931
θ=0.8	1.780400	0.0880289	1.707736	1.763652	0.17487

Table 1. Simulation results, $n = 100$

n=150	$\mu(\theta)$	c_1	m=$P_c^{L_2}$	$\bar{H}$	$sd_{H(\mu(\theta)\mid X_1,X_2,\ldots,X_n)}$
θ=0.2	0.445108	0.08798384	0.5012096	0.4436123	0.00149569
θ=0.4	0.890216	0.08798384	0.9060136	0.887553	0.00266295
θ=0.6	1.335300	0.08798384	1.3071587	1.326706	0.01095288
θ=0.8	1.780400	0.08798384	1.707462	1.76525	0.04263834

Table 2. Simulation results, $n = 150$

Simulation II

To proving the closeness of this new credibility premium, we make another simulation by Taking 9 values of θ with the following parameters: $h = 0.004$, $\omega = 0.9$, $\alpha = 2$ and $\beta = 8$, the same sample sizes are considered in this simulation.

n=100	$\mu(\theta)$	c_1	m=$P_c^{L_2}$	$\bar{H}$	$sd_{H(\mu(\theta)\mid X_1,X_2,\ldots,X_n)}$
θ=0.01	0.01002	0.059246202	0.021610547	0.01922807	0.00101804
θ=0.02	0.02004	0.059246202	0.030393611	0.02858928	0.008943467
θ=0.03	0.03006	0.059246202	0.040062203	0.03889427	0.008181747
θ=0.04	0.04008	0.059246202	0.048971771	0.04839034	0.009120648
θ=0.05	0.0501001	0.059246202	0.057808875	0.05780403	0.007527759
θ=0.06	0.06012012	0.059246202	0.066727051	0.06682108	0.001030524
θ=0.07	0.07014014	0.059246202	0.077531796	0.07676451	0.008697686
θ=0.08	0.08016016	0.059246202	0.083552218	0.08335358	0.00742907
θ=0.09	0.09036072	0.059246202	0.09637804	0.0947857	0.003017266

Table 3. Simulation results, $n = 100$

n=150	$\mu(\theta)$	c_1	m=P_c^{L2}	$\bar{H}$	$sd_{H(\mu(\theta)\|X_1,X_2,...,X_n)}$
θ=0.01	0.01004008	0.06906078	0.021616573	0.0187785	0.009932746
θ=0.02	0.02008016	0.06906078	0.030670718	0.02855261	0.008946811
θ=0.03	0.03012024	0.06906078	0.039429681	0.03800808	0.00718741
θ=0.04	0.04016032	0.06906078	0.048875386	0.04820519	0.008981859
θ=0.05	0.0502004	0.06906078	0.057242788	0.05723828	0.006831935
θ=0.06	0.06024048	0.06906078	0.066706573	0.06745428	0.00106605
θ=0.07	0.07028056	0.06906078	0.073971576	0.07529701	0.006379061
θ=0.08	0.08032064	0.06906078	0.084158238	0.08629355	0.008426812
θ=0.09	0.09036072	0.06906078	0.093061777	0.09590581	0.001535335

Table 4. Simulation results, $n = 150$

where, $\bar{H}(\mu(\theta)|X_1, X_2, ..., X_n)$ is denoted by $\bar{H}$.

Here, $\bar{H}(\mu(\theta)|X_1, X_2, ..., X_n)$ is the average of 1000 repetitions of $H(\mu(\theta)|X_1, X_2, ..., X_n)$, $sd_{H(\mu(\theta)|X_1,X_2,...,X_n)}$ denotes the standard deviation of $H(\mu(\theta)|X_1, X_2, ..., X_n)$, and $\mu(\theta)$ is the individual Esscher premium. The simulation shows better closeness of $\bar{H}(\mu(\theta)|X_1, X_2, ..., X_n)$ to $\mu(\theta)$ than m. Moreover, the simulation results show that the new exact credibility premium $H(\mu(\theta)|X_1, X_2, ..., X_n)$ is much closed to the individual Esscher premium. Also, we observe if $\theta \to 1$, the new exact credibility premium $H(\mu(\theta)|X_1, X_2, ..., X_n)$ is more much closed to the individual Esscher premium. This closeness proves the consistency of this credibility premium.

5.1. Remark. In this work, we take only the single insurance contract, which is valid only when the collective premium is given.

6. Conclusion

This study investigated weighted balanced loss function under the Esscher principle and generalized credibility premiums. It then derived distribution-free credibility premiums. More precisely, it employed a distribution free approach under WBLF to obtain a simple and new credibility premium which it is a combination of the collective premium and the individual Esscher premium.

Using a numerical simulation approach, it obtained empirical premiums and illustrated whether the empirical premiums converged to the Esscher premium. Its numerical simulation provides evidence of the convergence of the new credibility premiums derived in the study, and shows a simple way to calculate credibility premiums for actuaries.

However, our study is limited in the case of Poisson-gamma in which the Esscher premium has a linear formula, but under the other combinations of the exponential family, the Esscher premium does not hold.

Another research topic should include the other cases of the exponential family to generalizing this study and giving to insurer the choice between distributions.

Acknowledgements

The authors are grateful for the comments and suggestions by the referee and the Editor. Their comments and suggestions greatly improved the article.

References

[1] Bühlmann, H. *Experience rating and credibility*, Astin Bulletin **4**, 199-207, 1967.

[2] Bühlmann, H., Gisler, A. *A Course in Credibility Theory and its Applications*, Springer, Netherlands, 2005.

[3] Dey, D.G., Ghosh, M., Strawderman, W. *On estimation with balanced loss functions*, Statistics and Probability Letters **45**, 97-101, 1999.

[4] Farsipour, N.S., Asgharzadhe, A. *Estimation of a normal mean relative to balanced loss functions*, Statistical Papers **45**, 279-286, 2004.

[5] Gómez Déniz. *On the use of the weighted balanced loss function to obtain credibility premiums*, International conference on mathematical and statistical modeling in Honor of Enrique Castillo, June 28-30, 2006.

[6] Gómez Déniz. *A generalization of the credibility theory obtained by using the weighted balanced loss function*, Insurance: Mathematics and Economics 42, 850-854, 2008.

[7] Goovaerts, M.J., De Vylder.F. and Haezendonck. J. *Insurance Premiums: Theory and Applications.* North-Holland. Amsterdam, 1984.

[8] Goovaerts, M.J. and Hoogstad, W.J. *Credibility Theory. In: Surveys of Actuarial Studies*, **4**, Nationale-Nederlanden N.V., Rotterdam, 1987.

[9] Gupta, S. and Berger, J. *Statistical Decision Theory and Related Topics*, Springer, New York, 371-390, 1994.

[10] Heilmann, W. *Decision theoretic foundations of credibility theory*, Insurance: Mathematics and Economics **8** (1), 77-95, 1989.

[11] Herzog, T.N. *Introduction to Credibility Theory*, 2nd edition, ACTEX Publications, Winsted, 1996.

[12] Jafari, M., Marchand, E., Parsian, A. *On estimation with weighted balanced-type loss function*, Statistics and Probability Letters **76**, 773-780, 2006.

[13] Kaas, R., Dannenburg, D., Goovaerts, M. *Exact credibility for weighted observations*, Astin Bulletin **27** (2), 287-295, 1996.

[14] Klugman, S.A., Panjer, H.H., Willmot, G.E. *Loss Models: From Data to Decisions*, John Wiley and Sons, New York, 2004.

[15] Maolin Pan, Rongming Wang, Xianyi Wu. *On the consistency of credibility premiums regarding Esscher principle*, Insurance: Mathematics and Economics **42**, 119-126, 2008.

[16] Najafabadi, Amir T. Payandeh. *A new approach to the credibility formula*, Insurance: Mathematics and Economics **46**, 334-338, 2010.

[17] Norberg, R. *Credibility theory. In: Encyclopedia of Actuarial Science*, Wiley, Chichester, UK, 2004.

[18] Rodermund, M. *Introduction, Foundations of Casualty Actuarial Science*, Chapter 1. New York: Casualty Actuarial Society, 1989.

[19] Van Heerwaarden. A.E., Kaas, R. and Goovaerts, M.J. *Optimal reinsurance in relation to ordering of risks*, Insurance: Mathematics and Economics **8**. (1), 11-17, 1989.

[20] Zellner, A. *Bayesian and non-Bayesian estimation using balanced loss function*, In: Gupta, S.S., Berger, J.O. (Eds.), *Statistical Decision Theory and Related Topics*. Springer, New York, pp. 371-390, 1994.