第 39 卷 第 2 期 江西师范大学学报(自然科学版) Vol. 39 No. 2
2015 年 3 月 Journal of Jiangxi Normal University(Natural Science) Mar. 2015

文章编号:1000-5862(2015)02-0177-12

The Robustness for Premium Calculations Using Bayesian Approaches

WU Xianyi

(Department of Statistics and Actuarial Science, East China Normal University, Shanghai 200241, China)

Abstract: Three tightly related problems regarding the premium calculation principles are considered. Firstly, Bayesian premiums are defined by using of Bayesian approaches associated with loss principles. Then, two problems regarding the robustness of premium calculation principles are investigated. One is the robustness of non-Bayesian premiums with respect to arbitrary contaminations. The other one is the robustness of Bayesian premiums with respect to the prior distributions by means of the ε-contamination arguments. Finally, the reaction of a premium with respect to the contaminations and the range of premium using the Esscher principle when the contamination distribution varies in a distribution class are discussed.

Key words: robustness; premium calculation principles; ε-contaminations; Bayesian approaches; Esscher premium principles; loss principles

中图分类号: F 840; F 224 **文献标志码:** A **DOI:** 10.16357/j.cnki.issn1000-5862.2015.02.13

0 Introduction

A premium calculation principle aims at settling a suitable premium associated with an insurance contract in the sense that it provides sufficient but not excessive compensation for the underlying claims, by which losses to the insurance company are caused. In practice, an individual premium is usually revised from time to time in accordance with the accumulating of its claim experience up to date to reflect the true risk level of the policy-holder so as to meet the requirements of the competitive insurance markets. This situation provides a platform where the use of Bayesian methods becomes very natural.

In fact, Bayesian methodology has been used in the insurance science widely in experiential rate making from the later 1960s when two remarkable papers by H. Bühlmann were published [1-2]. In his papers, the foundations for linear credibility are established by the least square argument. From then on, much interest has been concentrated on the applications of Bayesian theories to the credibility such that the theory of credibility has been extensively developed over the last forty years. For an introduction to credibility theory, see [3]. Recent accounts of Bayesian statistics in actuarial science can be found in S. Klugman [4], R. Kaas et al [5], Z. M. Landsman et al [6-8], U. E. Makov et al [9], J. A. Nelder et al [10], and V. R. Young [11-12], to name just a few, which emphasize in particular the Bayesian approaches to credibility. In the open literature, most of the applications of Bayesian methodology in principle contributed to the credibility theory to establish the conditions under which a Bayesian estimate/prediction can be reduced to a credible linear combination of claim experiences. While various premium calculation principles have been suggested in practice and in academic context, and Bayesian methodology has been extensively developed in statistical science, there has not yet been much work on the applications of Bayesian methodolegy which directly make the insurance premiums by Bayesian approaches in premium-making, especially when dealing with the experiential rate-

收稿日期:2014-12-10
基金项目:国家自然科学基金(71371074)和上海市哲学社会科学基金(2010BJB004)资助项目.
作者简介:吴贤毅(1966-),男,贵州大方人,教授,博士,博士生导师,主要从事精算学、统计推断以及随机调度的研究.

making.

One of the purposes of this article is to incorperate the Bayesian methods into the rate-making practice to settle a premium when the prior distribution is applicable in a pure Bayesian way. Among all others, the loss principle for premium calculations[13] is the one which would be best tackled under the Bayesian framework since it adopts the opinion of decision making. It has turned out in statistical science that it is fruitful to incorporate Bayesian approaches into the decision-making theory. Motivated by this observation, associated with an arbitrary loss function, we in this paper define the corresponding Bayesian premium and explore their properties. To be specific, we in this paper define two types of Bayesian Premiums, data-free premiums and data-dependent premiums, and develop a way which calculates the later by means of the former.

Moreover, just like in statistical practice, the idea of robustness is also of essential importance in actuarial science. It has been studied in, at latest, the later of 1980s by E. Kremer[14]. Later works include H. R. Künsch[15], E. Kremer[16], A. Gisler et al[17] and E. Gòmez-Déniz et. al[18], and so on. In the research of robustness in actuarial science, there are two lines followed by many authors in the reported works. One, rooted from E. Kremer[14], is to study the robustness of a premium principle on large claim size and the other, which includes most of the studies on robustness in actuarial science, focus on credible premiums to robustmize the credible premiums when large claims happen to the policies.

So, as another topic of this paper, we will also investigate two problems regarding the robustness of premium calculation principles. One is on the robust problem in the traditional sense, i. e. , the robustness of non-Bayesian premiums, which was examined by E. Kremer[14] for the first time. The problem addressed in his work is the robustness of premium principles with respect to claims of large size. The main idea is to contaminate the distribution, which is originally adopted to calculate the premium, with a degenerated distribution that concentrates its mass at a given point, by means of ε-contamination distribution classes. This idea can be incorporated with the ones commonly used in robustness in statistical science to motivate us to study the robustness with respect to arbitrary contaminations. The other issue we will discuss is the robustness of Bayesian premiums with respect to the prior distribution (or construction function, in actuarial terminologies) under the framework of ε-contamination. Two aspects are discussed for the later issue. One is the reaction of the premium with respect to the contaminations. The other examines the boundaries of the premiums when the contamination distribution varies in a class of unimodal distributions or class of symmetric unimodal distributions. Particularly, precise results are derived for Esscher Premium Principles. The previous research for this problem can be found in E. Gòmez-Déniz et al[18] in which the distribution is implicitly limited to Gamma-Poisson structure. Our contribution here is to release the Gamma-Poisson condition so as to extend the results of E. Gòmez-Déniz et al[18] to arbitrary Esscher principles.

Our main contributions in this paper include the followings.

(i) We define two types of the Bayesian premiums. One is on the data-free basis andthe other one is data-dependent, which gives an individual premium in accordance with his/her claim history. A method to calculate the data-dependent premium by means of the data-free premium is established.

(ii) For the non-Bayesian premiums, we derive their reactions to the contaminative distributions which extend the existed works. For the Bayesian premiums, the reactions to the contaminative prior distributions are presented to reflect the robustness of the premium with respect to the selection of prior distributions.

(iii) We also investigate the ranges of an Esscher premium under the Bayesian framework when its contaminative prior distribution varies over a collection of unimodal distributions, or a collection of unimodal and symetric distributions. These ranges measure the robustness of the Bayesian Esscher premiums in another direction.

The rest of the paper is structured as follows.

We provide in section 2 some preliminaries including boundaries of integral ratios and the representations of posteriors under ε-contaminations. Bayesian premiums are defined and the calculation formulae are demonstrated in section 3. In section 4, after the robustness of non-Bayesian premiums being examined, we deal with the robustness of Bayesian premiums by the reaction of the Bayesian premiums with respect to the ε-contaminations of the prior distributions. Section 5 is contributed to the discussion on the boundaries of the Bayesian premiums with respect to the ε-contaminations of the prior distributions. The article is concluded in section 6 with some remarks including further interesting problems.

1 Preliminaries

Some useful preliminaries are prepared in this section. We first address the boundaries of a ratio of two expectations when the associated distributions vary over the class of all distributions. After defining the ε-contamination class of distributions, we give the representations of the posteriors, which are essential when discussing the boundaries of premiums, in terms of the contaminated prior distributions. Some of the proofs for the lemmas are given for easy reference.

1.1 Boundaries for ratios of expectations

In the first place, we present a lemma on the boundaries of a type of integral ratios which will establish the building block for later discussion on the robustness.

Lemma 1 Denote by D a Borel subset of the real space R, Q the collection of all distributions on D, and let both $f(x)$ and $g(x)$, $g(x) > 0$, be two bounded functions on D. Then

$$\sup_{F\in Q}(\text{ or }\inf_{F\in Q})\frac{\int_D f(x)\,\mathrm{d}F(x)}{\int_D g(x)\,\mathrm{d}F(x)} = \sup_{x\in D}(\text{ or }\inf_{x\in D})\frac{f(x)}{g(x)}.$$

Proof Since for every $F \in Q$

$$\int_D f(x)\,\mathrm{d}F(x) = \int_D \frac{f(x)}{g(x)}g(x)\,\mathrm{d}F(x) \leqslant \sup_{x\in D}\frac{f(x)}{g(x)}\int_D g(x)\,\mathrm{d}F(x),$$

We have

$$\frac{\int_D f(x)\,\mathrm{d}F(x)}{\int_D g(x)\,\mathrm{d}F(x)} \leqslant \sup_{x\in D}\frac{f(x)}{g(x)}.$$

Namely,

$$\sup_{F\in Q}\frac{\int_D f(x)\,\mathrm{d}F(x)}{\int_D g(x)\,\mathrm{d}F(x)} \leqslant \sup_{x\in D}\frac{f(x)}{g(x)}. \tag{1}$$

On the other hand, fixing any $x \in D$ and considering the degenerated distribution function $F_x(t) = I_{(x,\infty)}(t)$ where I is the indicator function, it plainly follows that

$$\sup_{F\in Q}\frac{\int_D f(x)\,\mathrm{d}F(x)}{\int_D g(x)\,\mathrm{d}F(x)} \geqslant \frac{\int_D f(x)\,\mathrm{d}F_x(x)}{\int_D g(x)\,\mathrm{d}F_x(x)} = \frac{f(x)}{g(x)}.$$

That is,

$$\sup_{F\in Q}\frac{\int_D f(x)\,\mathrm{d}F(x)}{\int_D g(x)\,\mathrm{d}F(x)} \geqslant \sup_{x\in D}\frac{f(x)}{g(x)}. \tag{2}$$

The proof of the first inequality in the theorem is thus completed by (1) and (2). The inf part is proved similarly.

1.2 ε-Contamination Class

Let $\pi_0(\theta)$ be a fixed distribution on a measurable space (Θ, F_Θ). Let Q be the set of some probability distributions on (Θ, F_Θ). Define

$$\Gamma = \{\pi_\varepsilon(\theta) = (1-\varepsilon)\pi_0 + \varepsilon q : \varepsilon \in [0,1], q \in Q\}, \tag{3}$$

where $\pi_\varepsilon(\theta)$ is referred to as an ε-contamination of π_0 by q.

A natural choice of Q is $Q = \{$all distributions on $(\Theta, F_\Theta)\}$ suggest by P. J. Huber[19]. J. Berger et al[20] later argued that while convenient in mathematics, it is too large and must include many unreasonable distributions yielding such large a class of posterior distributions that the corresponding results will possesse less practical values. Particularly, when π_0 is a unimodal distribution, J. Berger et al[20] suggested that Q should only comprise all the unimodal distributions sharing the same mode with π_0. They thought that under the suggested form of Q, every π in Q will not only retain the same function features but also provide sufficiently many alternatives for the forms of, e. g., the density functions and the features of tails. And the range of the posterior distribution arising from the later one will smaller than that of P. J. Huber[19]. We will follow this line and let

$$Q=\{F: F \text{ is a unimodal distribution sharing the same mode with } \pi_0\} \tag{4}$$

if π_0 is unimodal, or

$$Q=\{F: F \text{ is a unimodal and symmetrical distribution sharing the same mode with } \pi_0\} \tag{5}$$

if π_0 is symmetric and unimodal. We denote by Q_1 and Q_2 the Q defined by (4) and (5), respectively.

The following lemma provides the representations of unimodal and unimodal symmetric density functions.

Lemma 2 (i) f is unimodal and symmetrical with mode θ_0 if and only if it can be represented as

$$f(x)=\int_0^\infty \frac{1}{2z} I_{[\theta_0-z,\theta_0+z]}(x)\,\mathrm{d}\Phi(z)$$

for some distribution $\Phi(z)$ on $[0,\infty)$.

(ii) f is unimodal with mode θ_0 if and only if it can be represented as

$$f(x)=\alpha\int_0^\infty \frac{1}{z} I_{[\theta_0-z,\theta_0]}(x)\,\mathrm{d}\Phi_1(z)+(1-\alpha)\int_0^\infty \frac{1}{z} I_{[\theta_0,\theta_0+z]}(x)\,\mathrm{d}\Phi_2(z)$$

for some distributions $\Phi_1(z)$ and $\Phi_2(z)$ on $[0,\infty)$ and $a=\int_0^{\theta_0} f(x)\,\mathrm{d}x\in[0,1]$.

This lemma follows from the representations of monotone functions, see [20].

We later need the following notations.

$$Q_{1U}(\theta_0)=\{F: F \text{ is a unimodal distribution on } R \text{ with midpoint } \theta_0\},$$

$$q(x)=\frac{\lambda}{z_1} I_{[\theta_0-z_1,\theta_0]}(x)+\frac{(1-\lambda)}{z_2} I_{[\theta_0,\theta_0+z_2]}(x),$$

$$z_1, z_2>0, \lambda\in[0,1], \tag{6}$$

$$Q_{2U}(\theta_0)=\{F: \text{the density } q(x) \text{ of } F \text{ is as in (6)}\},$$

and

$$Q_{3U}(\theta_0)=\{F: F \text{ is a unimodal distribution on a closed interval with } \theta_0 \text{ as an end (left or right)}\}.$$

Lemma 2 indicates that Q_1 and Q_2 are the hulls of Q_{1U} and Q_{2U} respectively.

1.3 Representations of Posteriors under ε-Contaminations

We now turn to the representations of posteriors under ε-contaminations. For this purpose, let $\pi_\varepsilon(\theta)$ be a distribution in Γ defined by (3) and, for brevity, we suppose that both $\pi_0(\theta)$ and $q(\theta)$ are density functions. Thus, under the condition that the prior is $\pi_\varepsilon(\theta)$, the marginal distribution of $X^{(n)}=(X_1, X_2\cdots, X_n)$ is

$$m(x^{(n)}\mid\pi_\varepsilon)=\int_\Theta f(x^{(n)}\mid\theta)\,\pi_\varepsilon\,\mathrm{d}\theta,$$

where $f(x^{(n)}\mid\theta)$ is the likelihood function of $X^{(n)}$. By the representation of $\pi_\varepsilon(\theta)$ (see (3)), we see that

$$m(x^{(n)}\mid\pi_\varepsilon)=(1-\varepsilon)\,m(x^{(n)}\mid\pi_0)+\varepsilon m(x^{(n)}\mid q).$$

That is, the marginal distribution of $X^{(n)}$ is still of the form of contaminations with $m(x\mid\pi_0)$ and $m(x\mid q)$ taking the places of π_0 and q in equation (3) respectively.

Furthermore, let

$$\gamma(x^{(n)})=\varepsilon m(x^{(n)}\mid q)/m(x^{(n)}\mid\pi_\varepsilon),$$

which is obviously independent of θ and relative to sample $x^{(n)}$. The posterior distribution of θ is

$$\pi_\varepsilon(\theta\mid x^{(n)})=\frac{f(x^{(n)}\mid\theta)\,\pi_\varepsilon(\theta)}{m(x^{(n)}\mid\pi_\varepsilon)}=(1-\gamma(x^{(n)}))\,\pi_0(\theta\mid x^{(n)})+\gamma(x^{(n)})\,q(\theta\mid x^{(n)}) \tag{7}$$

where $\pi_0(\theta\mid x^{(n)})$ and $q(\theta\mid x^{(n)})$ are the posterior distributions of θ with respect to priors $\pi_0(\theta)$ and $q(\theta)$, respectively. (7) shows that while the posterior possesses the same form as a contamination distribution, the contamination factor ε is now replaced by $\gamma(x^{(n)})$. Meanwhile, the roles of π_0 and q in equation (3) are played by $\pi_0(\theta\mid x^{(n)})$ and $q(\theta\mid x^{(n)})$, respectively.

We now present the following result regarding the conditional distribution of X given $X^{(n)}$, which is also essential in the discussion of robustness since it again represents the conditional distribution of X given $X^{(n)}$ as the contamination-like form.

Lemma 3 If the prior distribution is an ε-contamination as defined by (3), the conditional distribution of X given $X^{(n)}$ is

$$f_\varepsilon(x\mid x^{(n)})=(1-\gamma(x^{(n)}))f_{\pi_0}(x\mid x^{(n)})+\gamma(x^{(n)})f_q(x\mid x^{(n)}).$$

where

$$\varphi(x^{(n)})=\varepsilon m(x^{(n)}\mid q)/m(x^{(n)}\mid\pi_\varepsilon). \tag{8}$$

Proof First note that the distribution of X conditional on $X^{(n)}$ in terms of its density function is

$$f_\varepsilon(x \mid x^{(n)}) = f_\varepsilon(x, x^{(n)}) / m(x^{(n)} \mid \pi_\varepsilon) = \int_\Theta f(x, x^{(n)} \mid \theta)\pi_\varepsilon(\theta)\,\mathrm{d}\theta / \int_\Theta f(x^{(n)} \mid \theta)\pi_\varepsilon(\theta)\,\mathrm{d}\theta, \quad (9)$$

where f_ε means the corresponding marginal/conditional distributions in the case that prior is ε-contaminated. Similarly,

$$f_{\pi_0}(x \mid x^{(n)}) = \int_\Theta f(x, x^{(n)} \mid \theta)\pi_0(\theta)\,\mathrm{d}\theta / m(x^{(n)} \mid \pi_0),$$
$$f_q(x \mid x^{(n)}) = \int_\Theta f(x, x^{(n)} \mid \theta) q(\theta)\,\mathrm{d}\theta / m(x^{(n)} \mid q). \quad (10)$$

Substituting (3) into the numerator of the right hand side in (9), we have

$$f_\varepsilon(x \mid x^{(n)}) = [(1-\varepsilon)\int_\Theta f(x, x^{(n)} \mid \theta)\pi_0(\theta)\,\mathrm{d}\theta + \varepsilon\int_\Theta f(x, x^{(n)} \mid \theta) q(\theta)\,\mathrm{d}\theta] / [m(x^{(n)} \mid \pi_\varepsilon)].$$

It follows immediately from (10) that

$$f_\varepsilon(x \mid x^{(n)}) = [(1-\varepsilon) m(x^{(n)} \mid \pi_0) f_{\pi_0}(x \mid x^{(n)}) + \varepsilon m(x^{(n)} \mid q) f_q(x \mid x^{(n)})] / [m(x^{(n)} \mid \pi_\varepsilon)] = (1-\gamma(x^{(n)})) f_{\pi_0}(x \mid x^{(n)}) + \gamma(x^{(n)}) f_q(x \mid x^{(n)}).$$

The lemma is thus proved.

2 Bayesian Premiums via Decision-Making Theory

Let X be the underlying claim of an insurance contract for which the actuary is to make an adequate premium, denoted by $H[X]$, by means of the decision-making theory [13,21]. Under this theoretical framework, a loss function $L(x, \eta)$ is selected to measure the gap between the premium η charged and the true claim x such that the suitable premium $H[X]$ for X is determined by minimizing the expected loss $E[L(X, \eta)]$,

$$E[L(X, H[X])] = \min_\eta E[L(X, \eta)]. \quad (11)$$

In practice, the distribution of X can be supposed to be drawn from a family L of distributions indexed by a parameter/parameters θ such that different θ usually indicates different distributions. The set Θ composed by all possible values of θ is referred to as the parameter space. In such a setting, $H[X]$ determined by (11) must be relevant to the parameter θ. In fact, (11) should be rewritten as

$$E[L(X, H[X]) \mid \theta] = \min_\eta E[L(X, \eta) \mid \theta]. \quad (12)$$

The solution to (12) is denoted by $H[X \mid \theta]$ and known in literature as the individual premium or risk premium of X [18]. The parameter θ is generally used to identify insured individuals and thus is latent. Hence, $H[X \mid \theta]$ can't be used directly. However, one can assume a distribution, $\pi(\theta)$, say, for θ known as prior distribution in statistics and structure function in actuarial convention. Under this setting, one can obtain the parameter-free marginal distribution of X, $M(x \mid \pi)$ under which the solution to (11), independent of θ, is denoted by $H_\pi[X]$ and termed as the data-free Bayesian premium. Under a DFBP principle, the differences between individuals are erased and all policy holders are charged a same premium.

In a majority of real practice there is a claim experience $X^{(n)} = (X_1, X_2, \cdots, X_n)$ available. The individual premium is then adjusted by its claim experience to reflect the differences among every individual risk. Write X_{n+1} for the claim in the future period $n+1$. Let the joint cumulative distribution function of $(X^{(n)}; X)$ conditional on θ be specified by $F(x_1, x_2 \cdots x_n; x \mid \theta) \triangleq F(x^{(n)}; x \mid \theta)$ where $x^{(n)} = (x_1, x_2 \cdots x_n)$. Intuitively speaking, $X^{(n)}$ with a realization $x^{(n)}$ is the sample used to estimate the parameters θ, which, in turn, are used to infer the distribution of X so as to give the individual experiential premium corresponding to each individual risk level. Substituting the conditional distribution of X_{n+1} given $X^{(n)}$ into (11) for the distribution of X, we can obtain a premium that is represented by a measurable function of $X^{(n)}$, taking into account the prior $\pi(\theta)$, and thus denoted by $H_\pi[X_{n+1}, X^{(n)}]$. It can be easily seen that

$$E[L(X_{n+1}, H_\pi[X_{n+1}; X^{(n)}])] = \min_{\eta(X^{(n)})} E[L(X_{n+1}, \eta(X^{(n)}))] \quad (13)$$

where the minimization is taken over the class of all measurable functions $\eta(X^{(n)})$ (or nonnegative measurable functions when a negative premium is prohibited). Especially, it is well known that if the loss function is $L(x, \eta) = (x-\eta)^2$ and the minimization in (13) is taken over the class of linear combinations of $x^{(n)} = (x_1, x_2, \cdots, x_n)$ i.e., $a_0 + \sum a_i X_i$, then it leads to the well-known credible premium [2].

The following result is obvious:

1) $\min_{\eta(X^{(n)})} E[L(X_{n+1},\eta(X^{(n)}))] \leqslant \min_{\eta\in\mathbf{R}(\text{or }\mathbf{R}^+)} E[L(X_{n+1},\eta)]$

indicating the intuition that it is always advantageous to take use of the experiential data.

2) $H_\pi[X_{n+1};X^{(n)}] = H[X_{n+1}\mid X^{(n)}] = H_{\pi}(\theta\mid X^{(n)})[X_{n+1}\mid(X^{(n)},\theta)]$ (14)

where $H[X_{n+1}\mid X^{(n)}]$ indicates the premium calculated by (11) at the distribution of X conditional on $X^{(n)}$ and $H_{\pi(\theta\mid X^{(n)})}[X_{n+1}\mid(X^{(n)},\theta)]$ is the DFBP with the distributions of X and θ being replaced by the distributions of X_{n+1} conditional on $(X^{(n)},\theta)$ and θ conditional on $X^{(n)}$ respectively.

3) Suppose $X_1,\cdots,X_n,X_{n+1}$ are independent given θ, then $H_\pi[X_{n+1};X^{(n)}] = H_{\pi(\theta\mid X^{(n)})}[X_{n+1}]$.

4) In addition, if the prior is conjugate, denoted by $\pi(\theta,\delta)$ and $X_1,\cdots,X_n,X_{n+1}$ are independent and identically distributed given θ, such that $\pi(\theta\mid X^{(n)}) = \pi(\theta,\delta(X^{(n)}))$ for some measurable function $\delta(X^{(n)})$, then

$$H_\pi[X_{n+1};X^{(n)}] = H_{\pi(\theta,\delta(X^{(n)}))}[X_{n+1}].$$

Example 1 Bayesian Esscher Premiums. We now discuss the well-known Esscher principle to show how the Bayesian premium is computed. Suppose, $X_1,\cdots,X_n,X_{n+1}$ are supposed to be independent and identically distributed given θ.

A Esscher premium is defined by

$$H[X\mid\theta] = E[X\mathrm{e}^{\lambda X}\mid\theta]/E[\mathrm{e}^{\lambda X}\mid\theta]$$

where λ is a given constant for which $E[X\mathrm{e}^{\lambda X}\mid\theta]$ and $E[\mathrm{e}^{\lambda X}\mid\theta]$ are both finite[23-24]. It may be regarded as a loss principle under the loss function

$$L_\lambda(x,\eta) = (x-\eta)^2\mathrm{e}^{\lambda x}. \tag{15}$$

Thus, by iterated expectation, the data-free Bayesian Esscher premium is

$$H_\pi[X] = \frac{E[X\mathrm{e}^{\lambda X}]}{E[\mathrm{e}^{\lambda X}]} = \frac{E[H[X\mid\theta]E[\mathrm{e}^{\lambda X}\mid\theta]]}{E[E[\mathrm{e}^{\lambda X}\mid\theta]]}. \tag{16}$$

Moreover, the Bayesian premium for Esscher Principle (or Bayesian Esscher Premium) is calculated as, by equalities in (13),

$$H_\pi[X;X^{(n)}] = E[X\mathrm{e}^{\lambda X}\mid X^{(n)}]/E[\mathrm{e}^{\lambda X}\mid X^{(n)}] = E_{\pi(\theta\mid X^{(n)})}[E[X\mathrm{e}^{\lambda X}\mid\theta]]/E_{\pi(\theta\mid X^{(n)})}[E[\mathrm{e}^{\lambda X}\mid\theta]]. \tag{17}$$

Consider a contract, under which the claim follows a compound Poisson distribution as

$$X = \sum_{i=1}^{N} Y_i, \tag{18}$$

where N is a Poisson variable such that

$$\Pr(N=n\mid\theta) = \mathrm{e}^{-\theta}\theta^n/n!,\ n=0,1,\cdots. \tag{19}$$

Moreover, the claim sizes Y_i are independent muatually with a common distribution independent of θ and independent of θ. Write $\varphi(\lambda) = E\mathrm{e}^{\lambda Y_1}$ and $\varphi(\lambda) = EY_1\mathrm{e}^{\lambda Y_1}$. As extensively adopted in literature such as, for example, J. Eichenauer et al[25], S. Klugman[4], and E. Gòmez-Déniz[18], we suppose the Poisson-Gamma distribution structure. That is, θ is supposed to be drawn from a Gamma distribution $\Gamma(a,b)$ with super-parameters $a>0$ and $b>0$ and density function

$$\pi(\theta) = \frac{a^b}{\Gamma(b)}\theta^{b-1}\mathrm{e}^{-a\theta},\ \theta>0. \tag{20}$$

Suppose we calculate the premium by Esscher Principle with parameter λ. Simple algebraic computation shows that

$$E[\mathrm{e}^{\lambda X}\mid\theta] = \mathrm{e}^{\theta(\varphi(\lambda)-1)},$$

and

$$E[X\mathrm{e}^{\lambda X}\mid\theta] = \theta\varphi(\lambda)\,\mathrm{e}^{\theta(\varphi(\lambda)-1)},$$

and hence the true individual premium is

$$H[X\mid\theta] = E[X\mathrm{e}^{\lambda X}\mid\theta]/E[\mathrm{e}^{\lambda X}\mid\theta] = \theta\varphi(\lambda).$$

Under the Gamma-Poisson assumptions,

$$H_\pi[X] = b\varphi(\lambda)/[a+1-\varphi(\lambda)]. \tag{21}$$

Suppose that the experiential data of claim numbers in last n years is $N^{(n)} = (N_1,N_2,\cdots,N_n)$, which are, conditional on θ, independent and identically distributed as in (19). Consequently, the posterior distribution of θ is

$$\pi(\theta\mid N^{(n)}) \propto \mathrm{e}^{-(a+n)\theta}\theta^{\sum_{i=1}^{n}N_i+b-1} \sim \Gamma(a+n,\sum_{i=1}^{n}N_i+b).$$

So, replacing b and a in $H_\pi[X]$ in (21) with $a+n$ and $b+\sum_{i=1}^{n}N_i$ respectively, yields

$$H_\pi[X;X^{(n)}] = (b+\sum_{i=1}^{n}N_i)\varphi(\lambda)/[a+n+1-\varphi(\lambda)]. \tag{22}$$

3 Robustness vs the Class of ε-Contamination Distributions

We now turn to the robustness of premium calculation principles in this section. The study of the robustness of premium calculation on distributions is rooted from E. Kremer[14,25], who investigated the robustness on large claims. We here discussed is the framework of robustness with respect to an ε-contamination distributions family.

3.1 Robustness of Non-Bayesian Premiums

Definition 1 Suppose a premium principle H is well-defined on the class

$$\aleph_F = \{F(\varepsilon, G): F(\varepsilon, G) = (1-\varepsilon)F + \varepsilon G, \varepsilon \in [0,1]\},$$

where G is a fixed distribution. The reaction of H to distribution G at risk X (with distribution function F) is defined by

$$R_{H,F}(G) = \lim_{\varepsilon\to 0}(H[(1-\varepsilon)F + \varepsilon G] - H[F]/\varepsilon).$$

Let D be a class of distribution functions, and define $r_{H,D}(F) = \sup_{G\in D}|R_{H,F}(G)|$. Then we say that H is robust at X with respect to the distribution class D if $r_{H,D}(F) \neq \infty$.

Obviously, $R_{H,F}(G)$ reduces to $R_{H,F}(y)$ when G is a degenerated distribution concentrating its mass at y. We now give the main theorems of this section.

Theorem 1 If a premium principle can be expressed as the form $E[f(X)]$ with $f(\cdot)$ being some specified function, then $R_{H,F}(G) = H[G] - H[F]$.

Proof Since

$$H[F_\varepsilon(G)] = E_{F_\varepsilon(G)}[f(X)] = (1-\varepsilon)E_F[f(X)] + \varepsilon E_G[f(X)],$$

we see that

$$\frac{H[F_\varepsilon(G)] - E_F[f(X)]}{\varepsilon} = E_G[f(X)] - E_F[f(X)].$$

That is $R_{H,F}(G) = E_G[f(X)] - E_F[f(X)] = H[G] - H[F]$.

This theorem says that H is robust at X with respect to the distribution class D if and only if $\sup_{G\in D} H[G] \neq \infty$. Moreover, it should be noted by Theorem 1 that the ratio $(H[F_\varepsilon(G)] - E_F[f(X)])/\varepsilon$ is independent of ε. Namely, the rate of the variation of $H[F_\varepsilon(G)]$ is constant when ε varies.

Theorem 2 If a premium principle is a solution of a equation $E[u(X,y)] = C$ and the orders of limits and integrals involved are interchangeable, then

$$R_{H,F}(G) = \frac{C - E_G[u(X, H[F])]}{E_F[\partial u(X, H[F])/\partial y]},$$

where $\partial u(X, H[F])/\partial y$ is the partial derivative of $u(x,y)$ with respect to the second variable y at $x = X$ and $y = H[F]$.

Proof Obviously,

$$E_F[u(X, H[F])] = E_{F_\varepsilon(G)}[u(X, H[F_\varepsilon(G)])] = C.$$

That is

$$E_F[u(X, H[F])] = (1-\varepsilon)E_F[u(X, H[F_\varepsilon(G)])] + \varepsilon E_G[u(X, H[F_\varepsilon(G)])].$$

Rearranging the terms, we have

$$E_F[u(X, H[F_\varepsilon(G)])] - E_F[u(X, H[F])] = \varepsilon(E_F[u(X, H[F_\varepsilon(G)])] - E_G[u(X, H[F_\varepsilon(G)])]).$$

Using the mean-value theorem, we know that

$$(H[F_\varepsilon(G)] - H[F])E_F[\partial u(X,\xi)/\partial y] = \varepsilon(E_F[u(X, H[F_\varepsilon(G)])] - E_G[u(X, H[F_\varepsilon(G)])]),$$

where ξ is a quantity between $H[F]$ and $H[F_\varepsilon(G)]$.

Dividing the two sides by $\varepsilon E_F[\partial u(X,\xi)/\partial y]$ and letting ε trend to zero, we obtain

$$R_{H,F}(G) = \lim_{\varepsilon\to 0}(H[F_\varepsilon(G)] - H[F]/\varepsilon = \lim_{\varepsilon\to 0}\frac{(E_F[u(X, H[F_\varepsilon(G)])] - E_G[u(X, H[F_\varepsilon(G)])])}{E_F[\partial u(X,\xi)/\partial y]} = \frac{C - E_G[u(X, H[F])]}{E_F[\partial u(X, H[F])/\partial y]}.$$

The theorem is proved.

The followings are straightforward corollaries to the above two theorems.

Corollary 1 (i) (Expectation Principle) Given $H_\lambda[X] = (1+\lambda)E[X]$, the reaction of H is

$$R_{H,F}(G) = (1+\lambda)(E_G[X] - E_F[X]).$$

Particularly, if G is a distribution degenerated at y, (the case of large claim), then

$$R_{H,F}(G) = (1+\lambda)(y - E_F[X]).$$

(ii) (Zero-utility Principle) Given that $H[F]$ is the solution to $E[u(H[F] - X)] = 0$, the reaction of H is

$$R_{H,F}(G) = -E_G[u(H[F] - X)]/E_F[u'(H[F] - X)].$$

Though listed in the corollary are just two of many premium principles, we should note that the results in E. Kremer[14], besides (b) and (c), are special cases of Theorems 1 and 2.

3.2 Robustness of Bayesian Premiums

We now proceed with the robustness of premium principles to investigate the reaction of a Bayesian premium to its prior distribution. In the current circumstance, the reaction is still defined as in

Definition 1 with the places of F and G taken by π_0 and q in (3) respectively. The loss function $L(x,\eta)$ is supposed at this point to be unimodal and two times differentiable. The reaction of a DDBP to its prior is stated below.

Theorem 3 When the loss function $L(x,\eta)$ is unimodal and of continuously 2-ordered derivative, the reaction of the DDBP to its prior is

$$R_{H,\pi_0}(q) = -m(x^{(n)}\mid q)/m(x^{(n)}\mid\pi_0)\cdot\frac{E_{f_q}[\partial L(X,H_{\pi_0}[X;X^{(n)}])/\partial y]}{E_{f_{\pi_0}}[\partial^2 L(X,H_{\pi_0}[X;X^{(n)}])/\partial y^2]}$$

provided that the orders of integral operations are interchangeable.

Proof First short $f_{\pi_0}(x\mid X^{(n)})$, $f_q(x\mid X^{(n)})$ and $f_\varepsilon(x\mid X^{(n)})$ as f_{π_0}, f_q, and f_ε respectively and denote $g(x,y) = \partial L(x,y)/\partial y$. Obviously, under the conditions of the theorem, $H_{\pi_0}(X;X^{(n)})$ and $H_\varepsilon(X;X^{(n)})$ are the solutions to

$$E_{f_{\pi_0}}[g(X,H_{\pi_0}[X;X^{(n)}])] = E_{f_\varepsilon}[g(X,H_{\pi_\varepsilon}[X;X^{(n)}])] = 0.$$

That is, by Lemma 3, $E_{f_{\pi_0}}[g(X,H_{\pi_0}[X;X^{(n)}])] = (1-\varphi(x^{(n)}))E_{f_{\pi_0}}[g(X,H_\varepsilon[X;X^{(n)}])] + \varphi(x^{(n)})E_{f_q}[g(X,H_\varepsilon[X;X^{(n)}])]$. Rearranging the terms, we obtain $E_{f_{\pi_0}}[g(X,H_\varepsilon[X;X^{(n)}])] - E_{f_{\pi_0}}[g(X,H_{\pi_0}[X;X^{(n)}])] = \varphi(x^{(n)})(E_{f_{\pi_0}}[g(X,H_\varepsilon[X;X^{(n)}])] - E_{f_q}[g(X,H_\varepsilon[X;X^{(n)}])])$.

Using the mean-value theorem, we know that $(H_{\pi_\varepsilon}[X;X^{(n)}] - H_{\pi_0}[X;X^{(n)}])E_{f_{\pi_0}}[\partial g(X,\xi)/\partial y] = \varphi(x^{(n)})(E_{f_{\pi_0}}[g(X,H_{\pi_\varepsilon}[X;X^{(n)}])] - E_{f_q}[g(X,H_{\pi_\varepsilon}[X;X^{(n)}])])$,

where ξ is a quantity in between $H_{\pi_0}[X;X^{(n)}]$ and $H_{\pi_\varepsilon}[X;X^{(n)}]$. Dividing the two sides with $\varepsilon E_{f_{\pi_0}}[\partial g(X,\xi)/\partial y]$ and letting ε trend to zero yields

$$R_{H,\pi_0}(q) = \lim_{\varepsilon\to 0}(H_{\pi_\varepsilon}[X;X^{(n)}] - H_{\pi_0}[X;X^{(n)}]/\varepsilon) = \lim_{\varepsilon\to 0}\varphi(x^{(n)})\,[E_{f_{\pi_0}}[g(X,H_{\pi_\varepsilon}[X;X^{(n)}])]/\varepsilon - E_{f_q}[g(X,H_{\pi_\varepsilon}[X;X^{(n)}])]]/E_{f_{\pi_0}}[\partial g(X,\xi)/\partial y]. \quad (23)$$

By formula (8)

$$\lim_{\varepsilon\to 0}\frac{\varphi(x^{(n)})}{\varepsilon} = \lim_{\varepsilon\to 0}\frac{m(x^{(n)}\mid q)}{m(x^{(n)}\mid\pi_\varepsilon)} = \frac{m(x^{(n)}\mid q)}{m(x^{(n)}\mid\pi_0)}. \quad (24)$$

Besides, since

$$\lim_{\varepsilon\to 0}E_{f_{\pi_0}}[\frac{\partial}{\partial y}g(X,\xi)] = E_{f_{\pi_0}}[\frac{\partial}{\partial y}g(X,H_{\pi_0}[X;X^{(n)}])],$$

$$\lim_{\varepsilon\to 0}E_{f_{\pi_0}}[g(X,H_{\pi_\varepsilon}[X;X^{(n)}])] = E_{f_{\pi_0}}[g(X,H_{\pi_0}[X;X^{(n)}])] = 0$$

and

$$\lim_{\varepsilon\to 0}E_{f_q}[g(X,H_{\pi_\varepsilon}[X;X^{(n)}])] = E_{f_q}[g(X,H_{\pi_0}[X;X^{(n)}])],$$

it obviously follows by considering (23) and (24) that

$$R_{H,\pi_0}(q) = -m(x^{(n)}\mid q)/m(x^{(n)}\mid\pi_\varepsilon)\cdot E_{f_q}[\partial L(X,H_{\pi_0}[X;X^{(n)}])/\partial y]/E_{f_{\pi_0}}[\partial^2 L(X,H_{\pi_0}[X;X^{(n)}])/\partial y^2],$$

which completes the proof.

We can also study the reaction of a DFBP to its prior distribution. Since the proof is similar to and simpler than the preceding one, the result is listed below without a proof.

Theorem 4 If the loss function $L(x,\eta)$ is unimodal and of continuously 2-ordered derivative, the reaction of the DFBP to its prior is

$$\widetilde{R}_{H,\pi_0}(q) = -\frac{E_{m(x\mid q)}[\partial L(X,H_{\pi_0}[X])/\partial y]}{E_{m(x\mid\pi_0)}[\partial^2 L(X,H_{\pi_0}[X])/\partial y^2]} \quad (25)$$

provided that the orders of integral operations are interchangeable with difference operations.

We conclude this section with an example for which we demonstrate the reaction of the Bayes Esscher principle to the prior.

Example 1 Reaction of the Bayesian Esscher principle to its prior. Since by (15)

$\partial L(x,y)/\partial y = 2(y-x)e^{\lambda x}$, $\partial^2 L(x,y)/\partial y^2 = 2e^{\lambda x}$,

we obtain by Theorem 4 that

$$\widetilde{R}_{H,\pi_0}(q) = -\frac{E_{m(x\mid q)}[\partial L(X,H_{\pi_0}[X])/\partial y]}{E_{m(x\mid\pi_0)}[\partial^2 L(X,H_{\pi_0}[X])/\partial y^2]} = \frac{E_{m(x\mid q)}[(H_{\pi_0}[X]-X)e^{\lambda X}]}{E_{m(x\mid\pi_0)}[e^{\lambda X}]}.$$

Recalling the relations

$$H_{\pi_0}[X] = E_{m(x\mid\pi_0)}[Xe^{\lambda X}]/E_{m(x\mid\pi_0)}[e^{\lambda X}],$$

$$H_q[X] = E_{m(x\mid q)}[Xe^{\lambda X}]/E_{m(x\mid q)}[e^{\lambda X}],$$

it follows by (25) that

$$\widetilde{R}_{H,\pi_0}(q) = -\frac{E_{m(x\mid q)}[e^{\lambda X}]}{E_{m(x\mid\pi_0)}[e^{\lambda X}]}(H_q[X] - H_{\pi_0}[X]). \quad (26)$$

Similarly, by Theorem3,

$$R_{H,\pi_0}(q) = -\frac{m(x^{(n)}\mid q)}{m(x^{(n)}\mid\pi_0)}\frac{E_{f_q}[\frac{\partial}{\partial y}L(X,H_{\pi_0}[X;X^{(n)}])]}{E_{f_{\pi_0}}[\frac{\partial^2}{\partial y^2}L(X,H_{\pi_0}[X;X^{(n)}])]} =$$

$$-\frac{m(x^{(n)}\mid q)}{m(x^{(n)}\mid \pi_0)}\frac{E_{f_q}[H_{\pi_0}[X;X^{(n)}]-X)\mathrm{e}^{\lambda X})]}{E_{f_{\pi_0}}[\mathrm{e}^{\lambda X}]}.$$

That is,

$$R_{H,\pi_0}(q)=-\frac{m(x^{(n)}\mid q)}{m(x^{(n)}\mid \pi_0)}\frac{E_{f_q}[\mathrm{e}^{\lambda X}]}{E_{f_{\pi_0}}[\mathrm{e}^{\lambda X}]}\cdot (H_{\pi_0}[X;X^{(n)}]-H_q[X;X^{(n)}]). \quad (27)$$

The reactions of DDBP and DFBP for Esscher principle are computed by (26) and (27) respectively.

4 Boundaries of Bayesian Esscher Principles

In Bayesian statistics, the robustness of an approach is also measured, among others, by the range of the statistic when the prior varies over some selected ε-contamination class [20]. In this section, we follow this line to study the range of the premiums calculated by a Bayesian principle, with or without a claim data set, when the prior varies over the ε-contamination class. As argued before in section 1.2, we discuss two types of contamination classes. One is the class of unimodal distributions and the other the class of unimodal symmetric distributions. The analyses is demonstrated under Esscher principle due to its convenient form of ratio of expectations, which allows for us to make use of traditional tools to derive our results. This problem has been studied in E. Gòmez-Déniz et al [18] in which the distribution is implicitly limited to Gamma-Poisson structure. Our results here extend those of E. Gòmez-Déniz et al [18] to arbitrary Esscher principles.

For simplicity, we suppose that $\Theta=\mathbf{R}$ in this section so as to avoid technical trivial details. In addition, we let

$$\Gamma_i=\{\pi_\varepsilon(\theta)=(1-\varepsilon)\pi_0(\theta)+\varepsilon q(\theta):\ \varepsilon\in[0,1],q\in Q_i\},i=1,2,$$

where Q_1 and Q_2 are defined as in (4) and (5), respectively.

The main results are stated in theorems below. The first two are on the DFBPs while the last two for DDBPS.

Theorem 2 (DFBP) The supremums and infimums of a DFBP are

(i) $\sup_{\pi_\varepsilon\in\Gamma_1}H_{\pi_\varepsilon}[X]=\sup_{q\in Q_{1U}}H_{\pi_\varepsilon}[X]$,
$\sup_{\pi_\varepsilon\in\Gamma_2}H_{\pi_\varepsilon}[X]=\sup_{q\in Q_{2U}}H_{\pi_\varepsilon}[X]$.

(ii) $\inf_{\pi_\varepsilon\in\Gamma_1}H_{\pi_\varepsilon}[X]=\inf_{q\in Q_{1U}}H_{\pi_\varepsilon}[X]$,
$\inf_{\pi_\varepsilon\in\Gamma_2}H_{\pi_\varepsilon}[X]=\inf_{q\in Q_{2U}}H_{\pi_\varepsilon}[X]$.

Proof (i) Let $q\in Q_1$. By definition of $H_{\pi_\varepsilon}[X]$ (see (16)),

$$H_{\pi_\varepsilon}[X]=\frac{(1-\varepsilon)E_{m(x\mid\pi_0)}[X\mathrm{e}^{\lambda X}]+\varepsilon E_q[E[X\mathrm{e}^{\lambda X}\mid\theta]]}{(1-\varepsilon)E_{m(x\mid\pi_0)}[\mathrm{e}^{\lambda X}]+\varepsilon E_q[E[\mathrm{e}^{\lambda X}\mid\theta]]}.$$

We at this point denote $g(\theta)=E(X\mathrm{e}^{\lambda X}\mid\theta)$ and $h(\theta)=E(\mathrm{e}^{\lambda X}\mid\theta)$, then $H_{\pi_\varepsilon}[X]$ is expressed as

$$H_{\pi_\varepsilon}[X]=\frac{\int_{-\infty}^{\infty}[(1-\varepsilon)E_{m(x\mid\pi_0)}[X\mathrm{e}^{\lambda X}]+\varepsilon g(\theta)]q(\theta)\,\mathrm{d}\theta}{\int_{-\infty}^{\infty}[(1-\varepsilon)E_{m(x\mid\pi_0)}[\mathrm{e}^{\lambda X}]+\varepsilon h(\theta)]q(\theta)\,\mathrm{d}\theta}. \quad (28)$$

In view of the representation of $q(\theta)$ in Lemma 2, we obtain

$$\begin{aligned}H_{\pi_\varepsilon}[X]&=\Big(\int_{-\infty}^{\infty}[(1-\varepsilon)E_{m(x\mid\pi_0)}[X\mathrm{e}^{\lambda X}]+\varepsilon g(\theta)]\cdot\int_0^{\infty}\frac{1}{2z}I_{[\theta_0-z,\theta_0+z]}(\theta)\,\mathrm{d}F(z)\,\mathrm{d}\theta\Big)\Big/\Big(\int_{-\infty}^{\infty}[(1-\varepsilon)\cdot E_{m(x\mid\pi_0)}[\mathrm{e}^{\lambda X}]+\varepsilon h(\theta)]\int_0^{\infty}\frac{1}{2z}I_{[\theta_0-z,\theta_0+z]}(\theta)\,\mathrm{d}F(z)\,\mathrm{d}\theta\Big)\\&=\Big(\int_0^{\infty}\frac{1}{2z}\int_{-\infty}^{\infty}[(1-\varepsilon)E_{m(x\mid\pi_0)}[X\mathrm{e}^{\lambda X}]+\varepsilon g(\theta)]\cdot I_{[\theta_0-z,\theta_0+z]}(\theta)\,\mathrm{d}\theta\mathrm{d}F(z)\Big)\Big/\Big(\int_0^{\infty}\frac{1}{2z}\int_{-\infty}^{\infty}[(1-\varepsilon)\cdot E_{m(x\mid\pi_0)}[\mathrm{e}^{\lambda X}]+\varepsilon h(\theta)]I_{[\theta_0-z,\theta_0+z]}(\theta)\,\mathrm{d}\theta\mathrm{d}F(z)\Big).\end{aligned}$$

Therefore, by Lemma 1,

$$\begin{aligned}\sup_{q\in Q_1}H_{\pi_\varepsilon}[X]&=\sup_{z>0}\Big(\int_{-\infty}^{\infty}[(1-\varepsilon)E_{m(x\mid\pi_0)}[X\mathrm{e}^{\lambda X}]+\frac{\varepsilon}{2z}g(\theta)]I_{[\theta_0-z,\theta_0+z]}(\theta)\,\mathrm{d}\theta\Big)\Big/\Big(\int_{-\infty}^{\infty}[(1-\varepsilon)\cdot E_{m(x\mid\pi_0)}[\mathrm{e}^{\lambda X}]+\frac{\varepsilon}{2z}h(\theta)]I_{[\theta_0-z,\theta_0+z]}(\theta)\,\mathrm{d}\theta\Big)\\&=\sup_{z>0}\Big((1-\varepsilon)E_{m(x\mid\pi_0)}[X\mathrm{e}^{\lambda X}]+\frac{\varepsilon}{2z}\cdot\int_{-\infty}^{\infty}g(\theta)\cdot I_{[\theta_0-z,\theta_0+z]}(\theta)\,\mathrm{d}\theta\Big)\Big/\Big((1-\varepsilon)E_{m(x\mid\pi_0)}[\mathrm{e}^{\lambda X}]+\frac{\varepsilon}{2z}\int_{-\infty}^{\infty}h(\theta)I_{[\theta_0-z,\theta_0+z]}(\theta)\,\mathrm{d}\theta\Big)=\sup_{q\in Q_{1U}}H_{\pi_\varepsilon}[X].\end{aligned}$$

Thus the first equality of (i) is proved. The rest of the theorem can be shown similarly.

For the case of unimodal prior distribution, futher results are available which is listed in the following.

Theorem 3 when $\pi_\varepsilon \in \Gamma_2$, then the supremum and infimum of the Data-Free Bayesian Premiums are $\sup_{\pi_\varepsilon \in \Gamma_2} H_{\pi_\varepsilon}[X] = \sup_{q \in Q_{3U}} H_{\pi_\varepsilon}[X]$ and $\inf_{\pi_\varepsilon \in \Gamma_2} H_{\pi_\varepsilon}[X] = \inf_{q \in Q_{3U}} H_{\pi_\varepsilon}[X]$.

Proof It can be easily proved by taking supremum first with respect to α, the mixing factor in representation of unimodal $q \in Q_{2U}$ (see(6)).

The similar theorems to the two stated above are present in the following on boundaries of a DDBP.

Theorem 4(DDBP) Suppose that $X_1, \cdots, X_n, X$ are independent and identically distributed conditional on θ. Then

(i) the supremum of the Bayesian premiums is

$$\sup_{q \in Q_1} H_{\pi_\varepsilon}[X, X^{(n)}] = \sup_{q \in Q_{1U}} H_{\pi_\varepsilon}[X, X^{(n)}],$$

$$\sup_{q \in Q_2} H_{\pi_\varepsilon}[X, X^{(n)}] = \sup_{q \in Q_{2U}} H_{\pi_\varepsilon}[X, X^{(n)}]. \quad (29)$$

(ii) (i) holds when sup is replaced by inf.

Proof We again denote $g(\theta) = E[Xe^{\lambda X} \mid \theta]$ and $h(\theta) = E[e^{\lambda X} \mid \theta]$. By the representations of $H_{\pi_\varepsilon}[X, X^{(n)}]$ and $\pi(\theta \mid X^{(n)})$, see (17) and (7) respectively, we have

$$H_{\pi_\varepsilon}[X, X^{(n)}] = H_{\pi(\theta \mid X^{(n)})}[X] = E_{\pi(\theta \mid X^{(n)})}[E[Xe^{\lambda X} \mid \theta]] / E_{\pi(\theta \mid X^{(n)})}[E[e^{\lambda X} \mid \theta]] = ((1-\gamma(X^{(n)}))E_{\pi_0(\theta \mid X^{(n)})}[g(\theta)] + \gamma(X^{(n)})E_{q(\theta \mid X^{(n)})}[g(\theta)]) / ((1-\gamma(X^{(n)})) \cdot E_{\pi_0(\theta \mid X^{(n)})}[h(\theta)] + \gamma(X^{(n)})E_{q(\theta \mid X^{(n)})}[h(\theta)]).$$

Letting

$$A = (1-\gamma(X^{(n)}))E_{\pi_0}[Xe^{\lambda X} \mid X^{(n)}],$$

$$B = (1-\gamma(X^{(n)}))E_{\pi_0}[e^{\lambda X} \mid X^{(n)}],$$

it follows that

$$H_{\pi_\varepsilon}[X, X^{(n)}] = \frac{A + \gamma(X^{(n)})E_{q(\theta \mid X^{(n)})}[g(\theta)]}{B + \gamma(X^{(n)})E_{q(\theta \mid X^{(n)})}[h(\theta)]}.$$

Clearly,

$$E_{q(\theta \mid X^{(n)})}[g(\theta)] = \int_\Theta g(\theta) f(X^{(n)}, \theta)\,d\theta \Big/ \int_\Theta f(X^{(n)}, \theta)\,d\theta = \int_\Theta g(\theta) f(X^{(n)} \mid \theta) q(\theta)\,d\theta \Big/ \int_\Theta f(X^{(n)}, \theta)\,d\theta.$$

Similarly,

$$E_{q(\theta \mid X^{(n)})}[h(\theta)] = \frac{\int_\Theta h(\theta) f(X^{(n)} \mid \theta) q(\theta)\,d\theta}{\int_\Theta f(X^{(n)}, \theta)\,d\theta}.$$

Consequently,

$$H_{\pi_\varepsilon}[X, X^{(n)}] = \left(A\int_\Theta f(X^{(n)}, \theta)\,d\theta + \gamma(X^{(n)})\int_\Theta g(\theta) f(X^{(n)} \mid \theta) q(\theta)\,d\theta\right) \Big/ \left(B\int_\Theta f(X^{(n)}, \theta)\,d\theta + \gamma(X^{(n)})\int_\Theta h(\theta) f(X^{(n)} \mid \theta) q(\theta)\,d\theta\right). \quad (30)$$

Therefore, analogizing (30) to (28) with $g(\theta) f(X^{(n)} \mid \theta)$ and $h(\theta) f(X^{(n)} \mid \theta)$ in (30) taking the roles of $g(\theta)$ and $h(\theta)$ in (28) respectively, we obtain

$$\sup_{q \in Q_1} H_{\pi_\varepsilon}[X, X^{(n)}] = \sup_{z>0}[A + \gamma(X^{(n)})\int_\Theta g(\theta) f(X^{(n)} \mid \theta) \cdot I_{[\theta_0 - z, \theta_0 + z]}(\theta)\,d\theta/2z] / [B + \gamma(X^{(n)})\int_\Theta h(\theta) f(X^{(n)} \mid \theta) \cdot I_{[\theta_0 - z, \theta_0 + z]}(\theta)\,d\theta/2z].$$

That is, $\sup_{q \in Q_1} H_{\pi_\varepsilon}[X, X^{(n)}] = \sup_{q \in Q_{1U}} H_{\pi_\varepsilon}[X, X^{(n)}]$.

So, we have proved the first equality of (29). The rest of (29) may be proved in the same way. (2) follows from the same arguments by replacing supremum with infimum correspondingly.

Following therem is a counterpart of Theorem 3 in the case of data-dependent premiums.

Theorem 5 When $\pi_\varepsilon \in \Gamma_2$, then the supremum and inmum of a DFBP are

$$\sup_{\pi_\varepsilon \in \Gamma_2} H_{\pi_\varepsilon}[X; X^{(n)}] = \sup_{q \in Q_{3U}} H_{\pi_\varepsilon}[X; X^{(n)}],$$

$$\inf_{\pi_\varepsilon \in \Gamma_2} H_{\pi_\varepsilon}[X; X^{(n)}] = \inf_{q \in Q_{3U}} H_{\pi_\varepsilon}[X; X^{(n)}].$$

Example 3 (The Example with Gamma-Poisson Structure) In order to illustrate the applications of the theorems, we end this section with an example regarding the Esscher principles with Gamma-Poisson distribution assumptions defined as in (18) ~ (20). The contamination is supposed to be unimodal with the same mode as $\pi_0(\theta)$, the distribution of which is prescribed by (20). That is, the mode for contamination $q(\theta)$ is $\theta_0 = (b-1)/a$ ($b > 1$ is implicitly assumed here). It can be regarded as a continuation to Example 1.

(i) (Data-Free Premium) Firstly, similar to(17), we have that

$$H_{\pi_\varepsilon}[X] = E_{\pi_\varepsilon}[E[Xe^{\lambda X} \mid \theta]] / E_{\pi_\varepsilon}[E[e^{\lambda X} \mid \theta]] = Ce^{\lambda C}E_{\pi_\varepsilon}[\theta e^{\theta(e^{\lambda C}-1)}] / E_{\pi_\varepsilon}[e^{\theta(e^{\lambda C}-1)}] \stackrel{\triangle}{=} Ce^{\lambda C}G,$$

where $G = E_{\pi_\varepsilon}[\theta e^{\theta(e^{\lambda C}-1)}] / E_{\pi_\varepsilon}[e^{\theta(e^{\lambda C}-1)}]$.

When $q(\theta) \in Q_{3U}(\theta_0)$,

$$G = E_{\pi_\varepsilon}[\theta e^{\theta(e^{\lambda C}-1)}] / E_{\pi_\varepsilon}[e^{\theta(e^{\lambda C}-1)}] = \frac{(1-\varepsilon)E_{\pi_0}[\theta e^{\theta(e^{\lambda C}-1)}] + \varepsilon E_q[\theta e^{\theta(e^{\lambda C}-1)}]}{(1-\varepsilon)E_{\pi_0}[e^{\theta(e^{\lambda C}-1)}] + \varepsilon E_q[e^{\theta(e^{\lambda C}-1)}]} = \frac{(1-\varepsilon)E_{\pi_0}[\theta e^{\theta(e^{\lambda C}-1)}]z + \varepsilon\int_{\theta_0}^{\theta_0+z}\theta e^{\theta(e^{\lambda C}-1)}\,d\theta}{(1-\varepsilon)E_{\pi_0}[e^{\theta(e^{\lambda C}-1)}]z + \varepsilon\int_{\theta_0}^{\theta_0+z} e^{\theta(e^{\lambda C}-1)}\,d\theta}.$$

Now that $E_{\pi_0}[\theta e^{\theta(e^{\lambda C}-1)}]$ and $E_{\pi_0}[\theta e^{\theta(e^{\lambda C}-1)}]$ have been computed respectively, the boundaries of $H_{\pi_\varepsilon}[X]$ is thus, by Theorem 3, represented as $Ce^{\lambda C}\inf_{z\in\mathbf{R}}G \leqslant H_{\pi_\varepsilon}[X] \leqslant Ce^{\lambda C}\sup_{z\in\mathbf{R}}G$, which can be numerically computed without any difficult for specified a, b, λ and C.

(ii) (Data-Dependent Premium) First, similar to (17), we have that

$$H_{\pi_\varepsilon}[X;X^{(n)}] = \frac{E_{\pi_\varepsilon(\theta\mid X^{(n)})}[E[Xe^{\lambda X}\mid\theta]]}{E_{\pi_\varepsilon(\theta\mid X^{(n)})}[E[e^{\lambda X}\mid\theta]]} = \frac{Ce^{\lambda C}E_{\pi_\varepsilon(\theta\mid X^{(n)})}[\theta e^{\theta(e^{\lambda C}-1)}]}{E_{\pi_\varepsilon(\theta\mid X^{(n)})}[e^{\theta(e^{\lambda C}-1)}]} \triangleq Ce^{\lambda C}G,$$

where $G = E_{\pi_\varepsilon(\theta\mid X^{(n)})}[\theta e^{\theta(e^{\lambda C}-1)}]/E_{\pi_\varepsilon(\theta\mid X^{(n)})}[e^{\theta(e^{\lambda C}-1)}]$.

Due to the representation formula (9),

$$G = E_{\pi_\varepsilon}[\theta e^{\theta(e^{\lambda C}-1)}]/E_{\pi_\varepsilon}[e^{\theta(e^{\lambda C}-1)}] = \Big((1-\gamma(x^{(n)}))\cdot E_{\pi_0(\theta\mid X^{(n)})}[\theta e^{\theta(e^{\lambda C}-1)}] + \gamma(x^{(n)})\cdot E_{q(\theta\mid X^{(n)})}[\theta e^{\theta(e^{\lambda C}-1)}]\Big)\Big/\Big((1-\gamma(x^{(n)}))\cdot E_{\pi_0(\theta\mid X^{(n)})}[e^{\theta(e^{\lambda C}-1)}] + \gamma(x^{(n)})E_{q(\theta\mid X^{(n)})}[e^{\theta(e^{\lambda C}-1)}]\Big). \quad (31)$$

For any distribution

$$q_z(\theta) = \begin{cases} I_{[\theta_0,\theta_0+z]}(\theta)/z, & z>0 \\ -I_{[\theta_0+z,\theta_0]}(\theta)/z, & z<0 \end{cases} \quad (32)$$

in Q_{3U} and the realization $(x_1,x_2,\cdots,x_n)$ of the historical claim numbers $N^{(n)}=(N_1,N_2,\cdots,N_n)$,

$$q(\theta\mid X^{(n)}) = \begin{cases} \frac{1}{z}e^{-n\theta}\theta\sum_{i=1}^{n}x_iI_{[\theta_0,\theta_0+z]}(\theta), & z>0, \\ -\frac{1}{z}e^{-n\theta}\theta\sum_{i=1}^{n}x_iI_{[\theta_0+z,\theta_0]}(\theta), & z<0. \end{cases} \quad (33)$$

$E_{q(\theta\mid X^{(n)})}[\theta e^{\theta(e^{\lambda C}-1)}]$ and $E_{q(\theta\mid X^{(n)})}[\theta e^{\theta(e^{\lambda C}-1)}]$ can be computed correspondingly using expression (33). Now that $E_{\pi_0(\theta\mid X^{(n)})}[\theta e^{\theta(e^{\lambda C}-1)}]$ and $E_{\pi_0(\theta\mid X^{(n)})}[e^{\theta(e^{\lambda C}-1)}]$ have been calculated before when deriving (22) the boundaries of $H_{\pi_\varepsilon}[X]$ is thus represented as

$$Ce^{\lambda C}\inf_z G \leqslant H_{\pi_\varepsilon}[X] \leqslant Ce^{\lambda C}\sup_z G \quad (34)$$

with G calculated by (31) ~ (33). Boundaries (34) can also be numerically computed.

5 Concluding Remarks

This paper has so far studied three tightly related problems involved in the premium calculations. One is the Bayesian approaches used in premium calculation principles based on the loss principle. The data-free Bayesian premiums are defined as the premium with respect to the parameter-free marginal distributions of claims. The algorithms for data-dependent premiums are generally developed to obtain the Bayes premiums by means of the data-free formulae. Esscher principle when it is understood as a loss principle is completely demonstrated as an example to show how a Bayesian premium is calculated.

The second one is on the robustness of Bayesian premiums. No matter a data set is used or not, the reactions of the premiums to the ε-contaminative distributions/priors are computed. The theory is presented in a general context. Some existed results are pointed to be straightforward corollaries of the results obtained here.

The last one is the boundaries of premiums in particular when the Esscher principle is used. That is, we have derived the boundaries for Esscher premiums when the contaminative priors vary over a class of unimodal distributions, or unimodal symmetrical distributions, which are of the same mode with the contaminated distribution.

The model for studying the reactions of premium principles to the contaminative distributions are of general features. However, as for the boundaries of the premiums, the model is not a general one [26]. So it would be an interesting future research direction to discuss the boundaries under general models. On the other hand, the contamination fashion adopted throughout this paper is only ε-contaminations. If other contamination models, such as, the one in M. Lavine [27] among others, are adopted, the corresponding conclusions are open now. So, it, in the authors' opinion, indicates another further interesting research topic.

6 References

[1] Bühlmann H. Experience rating and credibility [J]. ASTIN Bulletin, 1967, 4(3): 199-207.

[2] Bühlmann H. Experience rating and credibility [J]. ASTIN Bulletin, 1969, 5(2): 157-165.

[3] Herzog T N. An introduction to credibility [M]. 3rd ed. Winsted: ACTEX, 1999.

[4] Klugman S. Bayes statistics in actuarial science [M]. Bos-

ton: Kluwer Academic Publisher,1992.

[5] Kaas R,Dannenburg D,Goovaerts M. Exact credibility for weighted observations [J]. ASTIN Bulletin,1997,27(2):287-295.

[6] Landsman Z M,Makov U E. Exponential dispersion models and credibility [J]. Scandinavian Actuarial Journal,1998(1):89-96.

[7] Landsman Z M,Makov U E. On stochastic approximation and credibility [J]. Scandinavian Actuarial Journal,1999a(1):15-31.

[8] Landsman Z M,Makov U E. Sequential credibility evaluation for symmetric location claim distributions [J]. Insurance: Mathematics and Economics,1999b,24(3):291-300.

[9] Makov U E,Smith A F M,Liu Y H. Bayesian methods in actuarial science [J]. The Statistician,1996,45(4):503-515.

[10] Nelder J A,Verrall R J. Credibility theory and generalized linear models [J]. ASTIN Bulletin,1997,27(1):71-82.

[11] Young V R. Credibility using a loss function from spline theory [J]. Scandinavian Actuarial Journal,1997(2):160-185.

[12] Young V R. Credibility using a loss function from spline theory: parametric models with a one-dimensional sufficient statistic [J]. North American Actuarial Journal 1998,2(1):101-117.

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

[14] Kremer E. On robust premium principles [J]. Insurance: Mathematics and Economics,1986,5(4):271-274.

[15] Künsch H R. Robust methods for credibility [J]. ASTIN Bulletin,1992,22(1):33-49.

[16] Kremer E. Robust credibility via robust Kalman filtering [J]. ASTIN Bulletin,1994,24(2):221-233

[17] Gisler A,Reinhard P. Robust Credibility [J]. ASTIN Bulletin,1993,23(1):117-143.

[18] Gòmez-Déniz E,Hernández-Bastida A,Vázque-Polo F J. The Esscher premium in risk theory: a Bayesian sensitivity study [J]. Insurance: Mathematics and Economics,1999,25(3):387-395.

[19] Huber P J. The use of Choquet capacities in statistics [J]. Bulletin of International Statistical Institute,1973,45(1):181-191.

[20] Berger J,Berliner L M. Robust Bayes and empirical Bayes analysis with ε-contaminated priors [J]. Annals of statistics,1986,14(2):461-486.

[21] Kamps U. On a class of premium principles including Esscher principles [J]. Scandinavia Actuarial Journal,1998(1):75-80.

[22] Bühlmann M,Straub E. Glaubwurdigkrit fur schadensatze [J]. Mitt Ver Schweizer Versicherungsmath,1970,70(1):111-133.

[23] Bühlmann H. An economic premium principle [J]. ASTIN Bulletin,1980,11(1):52-60.

[24] Bühlmann H. The General economic premium principle [J]. ASTIN Bulletin,1984,14(1):13-21.

[25] Eichenauer J,Lehn J,Retting S,et al. A gamma-minimax result in credibility theory [J]. Insurance: Mathematics and Economics,1988,7(1):49-57.

[26] Sivagansan S,Berger J D. Ranges of posterior measures for priors with unimodal contaminations [J]. The Annals of Statistics,1989,17(2):868-889.

[27] Lavine M. Sensitivity in Bayesian statistics: the prior and the likelihood [J]. Journal of American Statistical Association,1991,86(415):396-399.

基于贝叶斯方法的保费计算的稳健性质

吴贤毅

(华东师范大学金融与统计学院,上海 200241)

摘要: 探讨了 3 类关于保费计算原理的相关问题. 首先,结合贝叶斯方法和损失原理定义了贝叶斯保费; 然后,研究了 2 类保费计算原理的稳健性质问题: 带任意污染系数的非贝叶斯保费的稳健性质和基于 ε-污染方法讨论的贝叶斯保费关于先验分布的稳健性质; 最后,运用 Esscher 保费原理分析了当污染在某个分布类变化时保费对污染的响应以及保费的值域.

关键词: 稳健性质; 保费计算原理; ε-污染; 贝叶斯方法; Esscher 保费原理; 损失原理

(责任编辑: 曾剑锋)