STATISTICS IN TRANSITION new series, September 2021
Vol. 22, No. 3, pp. 123–140, DOI 10.21307/stattrans-2021-030
Received – 18.09.2020; accepted – 09.04.2021

Robust Bayesian insurance premium in a collective risk model with distorted priors under the generalised Bregman loss

Agata Boratyńska[1]

ABSTRACT

The article presents a collective risk model for the insurance claims. The objective is to estimate a premium, which is defined as a functional specified up to unknown parameters. For this purpose, the Bayesian methodology, which combines the prior knowledge about certain unknown parameters with the knowledge in the form of a random sample, has been adopted. The generalised Bregman loss function is considered. In effect, the results can be applied to numerous loss functions, including the square-error, LINEX, weighted square-error, Brown, entropy loss. Some uncertainty about a prior is assumed by a distorted band class of priors. The range of collective and Bayes premiums is calculated and posterior regret Γ-minimax premium as a robust procedure has been implemented. Two examples are provided to illustrate the issues considered - the first one with an unknown parameter of the Poisson distribution, and the second one with unknown parameters of distributions of the number and severity of claims.

Key words: classes of priors, posterior regret, distortion function, Bregman loss, insurance premium

1. Introduction

We consider a Bayesian collective risk model. Our objective is to estimate a premium, which is defined as a functional H that assigns to any risk S a real number $H(S)$, the premium for taking the risk S. In practical situations the premium $H(S)$ can be calculated if the distribution of the risk S is known. We shall consider the case in which the distribution of S or the premium $H(S)$ is specified up to an unknown parameter θ, thus the risk premium will be denoted by $H(\theta)$. The premium $H(\theta)$ can be calculated according to different principles, from the simplest net premium to more sophisticated ones (see Kaas et al. (2009), Furman and Zitikis (2008)). Next we ought to estimate $H(\theta)$. We will use the Bayesian methodology, which combines the prior knowledge about a parameter θ (defined by a prior distribution π) with the knowledge

[1] Warsaw School of Economics SGH, Collegium of Economic Analysis, Institute of Econometrics, Al. Niepodległości 162, 02-554 Warszawa, Poland. E-mail: aborata@sgh.waw.pl. ORCID: 0000-0001-7363-1960.

in the form of a random sample $X = (X_1, X_2, \dots, X_n)$, where the distribution of this random variable depends on θ. The quality of an estimator is measured by the expected value of a loss function. There are a lot of different loss functions considered in the literature (see Heilmann (1989), Gómez-Déniz (2008), Boratyńska (2008) and Karimnezhad and Parsian (2018) for more references). Its choice depends on the severity of the error related to overestimation or underestimation. The most popular square-error loss equally penalizes over- and under-estimation of the same magnitude, the LINEX loss with $c < 0$ gives a greater error for underestimation than for overestimation, under the generalized entropy loss an error depends on the ratio between the estimated function and a considered action (for definitions of losses see Table 1). Again under- and over-estimation are not penalized equally. We will use the generalized Bregman loss (GB loss) function introduced by Karimnezhad and Parsian (2018) (for definition see Section 2). The class of GB loss functions contains different losses (weighted, symmetric, asymmetric, precautionary). All the loss functions mentioned above belong to that class. Thus, a practitioner has the great family of loss functions and he can choose one that expresses the severity of the estimation error very well.

Now, having some prior information about a parameter $\theta \in \Theta$, described by a prior distribution π (we will use the same notation for a probability distribution and its density (p.d.f.) with respect to the chosen measure on a probability space Θ), and a loss function $L(H(\theta), a)$ (measuring an error between the estimated parameter $H(\theta)$ and our estimate a) we can calculate the collective premium $\widehat{H}_\pi^C$, which minimizes the expected loss

$$E_\pi L(H(\theta), a) = \int_\Theta L(H(\theta), a)\pi(d\theta)$$

in a class of actions $a \in R$.

If, additionally, we have a random sample $X = (X_1, X_2, \dots, X_n)$ and X has a p.d.f. depended on a parameter θ, then for every value x of a random variable X we can calculate a Bayes premium $\widehat{H}_\pi^B(x)$, which minimizes the posterior risk equal the expected value of the loss function, if θ has the posterior distribution, thus

$$R_x(\pi, a) = E_\pi(L(H(\theta), a)|x) = \int_\Theta L(H(\theta), a)\pi(d\theta|x),$$

where $\pi(\cdot\,|x)$ denotes the posterior p.d.f. and a denotes a chosen action. Two premiums (defined above) express two situations. For example, the first premium is a premium in a class of risk. The prior expresses the population behaviour of an unknown parameter θ. The second premium combines knowledge about the population and about one considered risk (a policy).

The collective and Bayes premiums depend on a choice of a prior. The elicitation of a prior is difficult and can be uncertain. To model uncertainty of the prior information the robust Bayesian inference uses a class Γ of priors. In literature there are a lot of different classes Γ of priors: parametric classes of priors, ε-contamination classes, density and distribution band classes, quantile classes. For general references see Berger (1994), Ríos Insua and Ruggeri (2000). In insurance the robust Bayesian analysis was considered in many papers, for example: Young (2000), Chan et al. (2008), Gómez-Déniz (2009), Karimnezhad and Parsian (2014), Boratyńska (2017). Most of them present parametric or ε-contamination classes. We will use a class of priors based on distortion functions defined by Arias-Nicolás et al. (2016) (for definition see Section 3). The class is easily elicited and interpretable. It is connected with the stochastic and likelihood ratio orders. It quantifies a prior uncertainty in terms of distortion of a cumulative distribution function (c.d.f.). A parametric class of priors very often has a fixed shape of a c.d.f. During elicitation of a prior a practitioner has only approximate knowledge about a prior and narrowing down to a certain parametric family may be unjustified. The family considered in the paper can be an alternative. In insurance this class was considered by Sánchez-Sánchez et al. (2019). The concept of distortion functions has been used in actuarial science to model risk measure (see, for example, Balbas et al. (2009)).

Having a class Γ of priors we choose a measure of robustness of a statistical procedure and some concept of optimality. As a measure of robustness the range of posterior quantity, like the Bayes estimator, can be considered. If the range is small, then one may used the Bayes estimator as the robust procedure with respect to misspecifications of the prior (see Berger (1994), Ríos Insua and Ruggeri (2000) and Arias-Nicolás et al. (2016), among others). On the other hand, if conclusions differ widely, we should aim at eliciting additional information about the prior. However, the expert may not be willing to provide more information, and the practitioner is interested in choosing a single action from the set of actions provided by a global procedure. In this moment we can choose several concepts of an optimal procedure: the stable procedure, conditional Γ-minimax procedure or posterior regret Γ-minimax (PRGM) procedure (see Sivaganesan and Berger (1989), Ríos Insua et al. (1995), Boratyńska (1997, 2002), Ríos Insua and Ruggeri (2000), among others). We will use the last concept. Given the imprecision in elicitation of a prior, we try to make a decision, and this decision cannot be a Bayes action for every prior in the class Γ. Thus, we choose an action (in our problem an estimator of a premium), which minimizes the maximum loss of optimality in the class Γ and the largest possible increase in risk, resulting from making the wrong choice of a prior distribution, is kept as small as possible. The PRGM estimator depends on bands of the Bayes estimator when a prior runs over the class Γ. Thus, computing a PRGM estimator is simple provided that we have procedures to compute the range of Bayes estimators.

The article is organized as follows. Section 2 presents a guide for collective, Bayes and PRGM premiums under the GB loss. Section 3 reviews the structure of the class of priors based on distortion functions. Considering the GB loss function we find the bands of the Bayes estimator for a distorted band class of priors, thus we can compute the PRGM estimators. We note that for every value of a random sample X the optimal PRGM premium is the Bayes premium with respect to one prior from the considered class of priors. Section 4 contains PRGM estimators of a premium in some actuarial models with the GB loss function. We present some generalization for the case in which an unknown parameter is bidimensional (it is a case where a parameter of a probability distribution of a number of claims and a parameter of a probability distribution of a severity of claims are unknown and some prior information about them is known). Section 5 contains some concluding remarks.

2. Collective, Bayes and PRGM premiums under the GB loss function

Generally, let X be an observed random variable with a p.d.f. $f(\cdot\,|\theta)$ indexed by a real unknown parameter θ. Suppose θ has a prior distribution π. Let $L(H(\theta), a)$ be the generalized Bregman loss function (GB loss), measuring the penalty of incorrect estimation of a premium $H(\theta)$ by a real decision action a, defined as follows:

$$L(H(\theta),a) = w(H(\theta))\left[\phi(g(a)) - \phi\left(g\big(H(\theta)\big)\right) - \left(g(a) - g\big(H(\theta)\big)\right)\phi'\left(g\big(H(\theta)\big)\right)\right],$$

where real functions w, g and ϕ are fixed and $w(H(\theta)) > 0$ for every value $H(\theta)$, $g(\cdot)$ is a monotone function and $\phi(\cdot)$ is a convex, differentiable function and $\phi'(g(\theta)) = \frac{d}{dz}\phi(z)|_{z=g(\theta)}$. The shape of the GB loss depends of the choice of functions w, g and ϕ, for example, taking $w(z) = e^{-cz}$, $g(z) = z$ and $\phi(z) = e^{cz}$ $(c \neq 0)$ we obtain the LINEX loss function introduced by Varian (1974), taking $w(z) = 1$, $g(z) = z$ and $\phi(z) = z^2$ we have the square-error loss. Table 1 presents some examples of the GB loss. The following theorem is the corollary of Theorem 3.1. in Karimnezhad and Parsian (2018).

Theorem 1. *Let $X = x$. Then, under the GB loss function and a prior π, the collective $\widehat{H}_\pi^C$ and Bayes $\widehat{H}_\pi^B(x)$ premiums satisfy the following equations:*

$$\phi'(g(\widehat{H}_\pi^C)) = \frac{E_\pi(w(H(\theta))\phi'(g(H(\theta))))}{E_\pi(w(H(\theta)))},$$

$$\phi'(g(\widehat{H}_\pi^B(x))) = \frac{E_\pi(w(H(\theta))\phi'(g(H(\theta)))|x)}{E_\pi(w(H(\theta))|x)}.$$

Now, suppose that our knowledge about a prior is described by a family Γ of priors. Let $r^C(\Gamma)$ and $r^B(\Gamma, x)$ denote the range of a collective and a Bayes premium when priors run the class Γ, respectively, thus if $X = x$, then

$$r^C(\Gamma) = \sup_{\pi\in\Gamma}\widehat{H}_\pi^C - \inf_{\pi\in\Gamma}\widehat{H}_\pi^C \qquad \text{and} \qquad r^B(\Gamma, x) = \sup_{\pi\in\Gamma}\widehat{H}_\pi^B(x) - \inf_{\pi\in\Gamma}\widehat{H}_\pi^B(x).$$

Consider the posterior regret of an action a given by

$$r_x(\pi, a) = R_x(\pi, a) - R_x(\pi, \widehat{H}_\pi^B(x)).$$

In a sense, for $X = x$, it measures the loss of optimality due to choosing a instead of the optimal Bayes estimate. The estimator $\widehat{H}_\Gamma^{PR}$ is the posterior regret Γ-minimax premium (PRGM premium) if for every value x of X

$$\inf_{a\in R}\sup_{\pi\in\Gamma} r_x(\pi, a) = \sup_{\pi\in\Gamma} r_x(\pi, \widehat{H}^{PR}(x)).$$

We will use the following theorem to calculate the PRGM premium.

Theorem 2. (Karimnezhad and Parsian (2018)) *In estimating $H(\theta)$ under the GB loss function, let $X = x$, Γ be a class of prior distributions and let $\underline{H} = \underline{H}(x) = \inf_{\pi\in\Gamma} \widehat{H}_\pi^B(x)$, $\overline{H} = \overline{H}(x) = \sup_{\pi\in\Gamma} \widehat{H}_\pi^B(x)$ and $\underline{H} < \overline{H}$.*
If $w(H) = const$, then

$$g(\widehat{H}^{PR}(x)) = \frac{\phi(g(\overline{H})) - \phi(g(\underline{H})) - (g(\overline{H})\phi'(g(\overline{H})) - g(\underline{H})\phi'(g(\underline{H})))}{\phi'(g(\underline{H})) - \phi'(g(\overline{H}))}.$$

If there exists a constant k such that $E_\pi(w(H(\theta))|x) = \frac{k}{\phi'(g(\widehat{H}_\pi^B(x)))}$, then

$$\frac{\phi(g(\widehat{H}^{PR}(x))) - \phi(g(\overline{H})) - \phi'(g(\overline{H}))(g(\widehat{H}^{PR}(x)) - g(\overline{H}))}{\phi(g(\widehat{H}^{PR}(x))) - \phi(g(\underline{H})) - \phi'(g(\underline{H}))(g(\widehat{H}^{PR}(x)) - g(\underline{H}))} = \frac{\phi'(g(\overline{H}))}{\phi'(g(\underline{H}))}.$$

Directly from the proof of Theorem 2 we have the following corollaries.

Corollary 1. *Under the assumptions of Theorem 2 for every x of X*

$$\underline{H}(x) \leq \widehat{H}^{PR}(x) \leq \overline{H}(x).$$

Corollary 2. *Under the assumptions of Theorem 2, if for every value x of X the set $\{\widehat{H}_\pi^B(x) : \pi \in \Gamma\}$ is a connected set, then for every x there exists $\pi \in \Gamma$ such that $\widehat{H}^{PR}(x) = \widehat{H}_\pi^B(x)$.*

Table 1 presents collective, Bayes and PRGM premiums for different loss functions belonging to the class of GB loss functions.

Table 1. Examples of GB loss functions and collective, Bayes and PRGM premiums (for more examples and details see Karimnezhad and Parsian (2018))

$L(H,a)$	$\hat{H}_\pi^C$	$\hat{H}_\pi^B$	$\hat{H}^{PR}$
square-error loss $(H-a)^2$	$E_\pi H$	$E_\pi(H\|x)$	$0.5(\underline{H}+\overline{H})$
LINEX loss $e^{c(a-H)}-c(a-H)-1$	$\frac{-1}{c}\ln E_\pi e^{-cH}$	$\frac{-1}{c}\ln E_\pi(e^{-cH}\|x)$	$\underline{H}+\frac{1}{c}\ln\left(\frac{c(\underline{H}-\overline{H})}{\exp(c(\underline{H}-\overline{H}))-1}\right)$
weighted squared loss (1) $\frac{1}{H}(a-H)^2$	$\left(E_\pi\frac{1}{H}\right)^{-1}$	$\left(E_\pi\frac{1}{H}\|x\right)^{-1}$	$\sqrt{\underline{H}\overline{H}}$
weighted squared loss (2) $\frac{1}{H^2}(a-H)^2$	$\frac{E_\pi\frac{1}{H}}{E_\pi\frac{1}{H^2}}$	$\frac{E_\pi(\frac{1}{H}\|x)}{E_\pi(\frac{1}{H^2}\|x)}$	Th.2 is not applicable
Brown loss $(\ln a-\ln H)^2$	$e^{E_\pi\ln H}$	$e^{E_\pi(\ln H\|x)}$	$\sqrt{\underline{H}\overline{H}}$
precautionary loss $\frac{H}{a}+\frac{a}{H}-2$	$\sqrt{\frac{E_\pi H}{E_\pi\frac{1}{H}}}$	$\sqrt{\frac{E_\pi(H\|x)}{E_\pi(\frac{1}{H}\|x)}}$	Th.2 is not applicable
generalized entropy loss $\left(\frac{a}{H}\right)^q-q\ln\frac{a}{H}-1$	$\left(E_\pi\left(\frac{1}{H^q}\right)\right)^{\frac{-1}{q}}$	$\left(E_\pi\left(\frac{1}{H^q}\|x\right)\right)^{\frac{-1}{q}}$	$\left(\frac{\ln\underline{H}^q-\ln\overline{H}^q}{\overline{H}^{-q}-\underline{H}^{-q}}\right)^{\frac{1}{q}}$

3. Distorted band class of priors

We start with recalling the definition of the stochastic and likelihood ratio orders and a distortion function.

Let π_1 and π_2 be two probability distributions on the space Θ and F_{π_1} and F_{π_2} their cumulative distribution functions. We say that π_1 is smaller than π_2 in the stochastic order (denoted by $\pi_1 \preceq \pi_2$) if and only if for every $t \in R$ we have $F_{\pi_1}(t) \geq F_{\pi_2}(t)$. We say that π_1 is smaller than π_2 in the likelihood ratio order (denoted by $\pi_1 \preceq_{lr} \pi_2$) if and only if the ratio of their densities $\frac{\pi_2(\theta)}{\pi_1(\theta)}$ increases over the union of the supports of π_1 and π_2 (here $a/0$ is taken to be equal to $+\infty$ whenever $a>0$ and a support of a p.d.f. π is a closure of a set $\{\theta \in \Theta: \pi(\theta)>0\}$).

Let V and W be two random variables such that $V \sim \pi_1$ and $W \sim \pi_2$. It is well known that

$$\pi_1 \preceq_{lr} \pi_2 \quad \Longrightarrow \quad \pi_1 \preceq \pi_2$$

and

$$\pi_1 \preceq \pi_2 \quad \Leftrightarrow \quad E\psi(V) \leq E\psi(W), \tag{$*$}$$

for all increasing functions ψ for which the expectations exist. For more details about stochastic orders see Shaked and Shanthikumar (2007), for the stochastic ordering of posterior distributions, marginal distributions of data and predictive distributions see Męczarski (2015).

Let $h: [0,1] \longrightarrow [0,1]$ be a nondecreasing, continuous function such that $h(0) = 0$ and $h(1) = 1$. Then h is called a distortion function. Let π be a probability distribution on Θ, then a probability distribution π_h, with a c.d.f. of the form $F_{\pi_h} = h(F_\pi)$, is called the distorted distribution with the distortion function h.

Suppose that the prior distribution is not exactly specified and consider the following class of priors.

Definition (Arias-Nicolás et al. (2016)). *Let $\bar{\pi}$ be a specific prior belief. The distorted band class $\Gamma_{\bar{\pi},h_1,h_2}$ associated with $\bar{\pi}$, based on h_1 and h_2, a concave and convex distortion functions, respectively, is defined as*

$$\Gamma_{\bar{\pi},h_1,h_2} = \{\pi\colon\ \bar{\pi}_{h_1} \preceq_{lr} \pi \preceq_{lr} \bar{\pi}_{h_2}\}.$$

The following properties are very useful (for details see Arias-Nicolás et al. (2016)):

- easy elicitation and structure,
- $\Gamma_{\bar{\pi},h_1,h_2} \subseteq \{\pi\colon\ \bar{\pi}_{h_1} \preceq \pi \preceq \bar{\pi}_{h_2}\}$,
- if $\pi_1, \pi_2 \in \Gamma_{\bar{\pi},h_1,h_2}$, then for every $\varepsilon \in [0,1]$ and $\pi_\varepsilon = (1-\varepsilon)\pi_1 + \varepsilon\pi_2$ we have $\pi_\varepsilon \in \Gamma_{\bar{\pi},h_1,h_2}$,
- for every $\pi \in \Gamma_{\bar{\pi},h_1,h_2}$ and every x the posterior distribution satisfies

$$\bar{\pi}_{h_1}(\cdot\,|x) \preceq_{lr} \pi(\cdot\,|x) \preceq_{lr} \bar{\pi}_{h_2}(\cdot\,|x).$$

Example 1. Let $\bar{\pi}$ be a fixed prior on the space Θ. Consider a class

$$\Gamma_1 = \{\pi\colon\ \bar{\pi}_{h_{1,c_1}} \preceq_{lr} \pi \preceq_{lr} \bar{\pi}_{h_{2,c_2}}\},$$

where h_{1,c_1}, h_{2,c_2} are two distortion functions such that

$$h_{1,c_1}(z) = 1-(1-z)^{c_1},\quad h_{2,c_2}(z) = z^{c_2},$$

$$\bar{\pi}_{h_{1,c_1}}(\theta) = \frac{d}{d\theta}(1-(1-F_{\bar{\pi}}(\theta))^{c_1}),\ \ \bar{\pi}_{h_{2,c_2}}(\theta) = \frac{d}{d\theta}((F_{\bar{\pi}}(\theta))^{c_2}),$$

and $c_1 > 1$, $c_2 > 1$ are fixed numbers. Thus, if c_1 and c_2 are integers, then the bounds distributions are the distributions of the first and the last order statistics. The following properties describe the dependence on parameters c_1 and c_2.

- If $c'_1 > c_1$, then $\bar{\pi}_{h_{1,c'_1}} \preceq_{lr} \bar{\pi}_{h_{1,c_1}}$.
- If $c'_2 > c_2$, then $\bar{\pi}_{h_{2,c_2}} \preceq_{lr} \bar{\pi}_{h_{2,c'_2}}$.
- Similar order is for posterior distributions.
- The Kolmogorov distance (see Arias-Nicolas et al. (2016))

$$dK\left(\bar{\pi},\bar{\pi}_{h_{1,c_1}}\right) = (c_1-1)c_1^{\frac{-c_1}{c_1-1}}\ ,\quad dK(\bar{\pi},\bar{\pi}_{h_{2,c_2}}) = (c_2-1)c_2^{\frac{-c_2}{c_2-1}}.$$

We will use that class for elicitation priors in Section 4.

Now, considering the GB loss we would like to find bounds of a set of Bayes estimators of the premium. The following lemma presents the preservation of order of the collective and Bayes premiums computed under the GB loss function when prior distributions are in the likelihood ratio order.

Lemma 1. *Let H be an increasing function of θ. Let π_1 and π_2 be two priors such that $\pi_1 \preceq_{lr} \pi_2$ and $\widehat{H}^C_{\pi_i}$, $\widehat{H}^B_{\pi_i}$ be the collective and Bayes premium under the GB loss and the prior π_i, for $i = 1, 2$. Then for every x of X*

$$\widehat{H}^C_{\pi_1} \leq \widehat{H}^C_{\pi_2} \quad and \quad \widehat{H}^B_{\pi_1}(x) \leq \widehat{H}^B_{\pi_2}(x).$$

If H is decreasing, then in the above inequalities there is the sign change.

Proof. Assume H is increasing (if H is decreasing, then the proof is similar, only we have opposite inequalities in $(**)$).

Having a probability distribution π and a positive integrable function w, define the probability distribution π^w with the p.d.f. equal $\pi^w(\theta) = \frac{w(H(\theta))\pi(\theta)}{\int_\Theta w(H(\theta))\pi(d\theta)}$. If $\pi_1 \preceq_{lr} \pi_2$, then $\frac{\pi_2^w(\theta)}{\pi_1^w(\theta)} = \frac{\int_\Theta w(H(\theta))\pi_1(d\theta)}{\int_\Theta w(H(\theta))\pi_2(d\theta)} \cdot \frac{\pi_2(\theta)}{\pi_1(\theta)}$ is an increasing function of θ, hence $\pi_1^w \preceq_{lr} \pi_2^w$ and $\pi_1^w \preceq \pi_2^w$.

Note that $\phi'(g(\widehat{H}^C_{\pi_i}))$ (see the formula in Theorem 1) is the expected value of the function $\phi'(g(H(\theta)))$ if θ has the probability distribution π_i^w, $i = 1,2$. Now, applying the property $(*)$ of the stochastic order, if g is increasing, we have

$$\phi'\left(g(\widehat{H}^C_{\pi_1})\right) \leq \phi'\left(g(\widehat{H}^C_{\pi_2})\right) \qquad (**)$$

(if g is decreasing we have opposite inequalities) and obtain the assertion for the collective premium. The proof for the Bayes premium is similar, we only put a posterior distribution $\pi(\cdot\,|x)$ in the place of π.

The following theorem presents the bounds of a set of Bayes estimators and it is a conclusion from Lemma 1.

Theorem 3. *Under the GB loss function and the distorted band class $\Gamma_{\overline{\pi},h_1,h_2}$ of priors, if H is an increasing function of θ and for every $\pi \in \Gamma_{\overline{\pi},h_1,h_2}$ and every x of X there exist $\widehat{H}^C_\pi$ and $\widehat{H}^B_\pi(x)$, then*

$$\inf_{\pi\in\Gamma_{\overline{\pi},h_1,h_2}} \widehat{H}^C_\pi = \widehat{H}^C_{\overline{\pi}_{h_1}}, \qquad \sup_{\pi\in\Gamma_{\overline{\pi},h_1,h_2}} \widehat{H}^C_\pi = \widehat{H}^C_{\overline{\pi}_{h_2}},$$
$$\inf_{\pi\in\Gamma_{\overline{\pi},h_1,h_2}} \widehat{H}^B_\pi = \widehat{H}^B_{\overline{\pi}_{h_1}}, \qquad \sup_{\pi\in\Gamma_{\overline{\pi},h_1,h_2}} \widehat{H}^B_\pi = \widehat{H}^B_{\overline{\pi}_{h_2}}.$$

If H is decreasing, then inf and sup change places.

Having the upper and lower bounds for the set of Bayes premiums and applying Theorem 2, we can calculate the PRGM premium if the class of priors is equal $\Gamma_{\overline{\pi},h_1,h_2}$.

Remarks

1. Arias-Nicolás et al. (2016) define the class of submodular loss functions and obtain the bounds of the set of Bayes actions under priors belonging to $\Gamma_{\overline{\pi},h_1,h_2}$, if a loss function is convex in a and submodular. If $w(\theta) = const$ then the GB loss is

submodular $\left(\frac{\partial^2 L(\theta,a)}{\partial\theta\partial a} = -g'(a)g'(\theta)\varphi''(g(\theta)) \leq 0\right)$, but if $w(\theta) \neq const$, then a GB loss may not have the submodularity property. As an example consider $L(\theta,a) = \frac{1}{\theta^2}(a-\theta)^2$.

2. Applying Remark 8 in Sánchez-Sánchez et al. (2019) and Corollaries 1 and 2 we obtain that for every x there exists $\pi_0 \in \Gamma_{\overline{\pi},h_1,h_2}$ such that $\widehat{H}^{PR}(x)$ is equal to the Bayes estimator with respect to the prior π_0.

Example 2. In that example we present the exact formula for the PRGM estimator for some GB losses and a certain class $\Gamma_{\overline{\pi},h_1,h_2}$ of priors.

Let X be an observed random variable with the negative binomial distribution, $bin^-(r,\theta)$, where $\theta \in (0,1)$ is unknown and $r > 0$ is known, with the p.d.f. given by $f(x|\theta) = \frac{\Gamma(r+x)}{\Gamma(r)x!}\theta^r(1-\theta)^x$, if $x = 0,1,2,\ldots$.

Let $\overline{\pi}$ be a prior of θ with the p.d.f. equal $\overline{\pi}(\theta) = 2\theta$ if $\theta \in (0,1)$. We are interested in estimating a function $H(\theta) = \frac{1-\theta}{\theta}$. Note that $E(X|\theta) = rH(\theta)$. Hence, if X describes the number of claims, then we are interested in estimating the expected value of the number of claims. Consider $h_1(z) = z^{0.75}$ and $h_2(z) = z^2$ and a class $\Gamma_{\overline{\pi},h_1,h_2}$ of priors. Then $F_{\overline{\pi}}(\theta) = \theta^2$, $F_{\overline{\pi}_{h_1}}(\theta) = \theta^{1.5}$, $F_{\overline{\pi}_{h_2}}(\theta) = \theta^4$ for $\theta \in (0,1)$. If $X = x$, then posterior distributions for priors $\overline{\pi}, \overline{\pi}_{h_1}$ and $\overline{\pi}_{h_2}$ are beta distributions $Beta(r+2, x+1)$, $Beta(r+0.5, x+1)$ and $Beta(r+4, x+1)$, where a beta distribution with parameters $\alpha > 0$ and $\beta > 0$, $Beta(\alpha,\beta)$, has the p.d.f. given by $\pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}$, if $\theta \in (0,1)$.

Table 2. Bayes and PRGM estimators and the oscillation $r^B(\Gamma_{\overline{\pi},h_1,h_2}, x)$ under some losses, notation: $A = \frac{\Gamma(r+1.5)}{\Gamma(r+1.5+q)}$ and $B = \frac{\Gamma(r+4)}{\Gamma(r+4+q)}$.

Loss function	$(H-a)^2$	$\frac{1}{H}(H-a)^2$	$\left(\frac{a}{H}\right)^q - q\ln\frac{a}{H} - 1$
$\widehat{H}^B_{\overline{\pi}}(x)$	$\frac{x+1}{r+1}$	$\frac{x}{r+2}$	$\left(\frac{\Gamma(r+2)x!}{\Gamma(r+2+q)\Gamma(x-q+1)}\right)^{\frac{1}{q}}$
$\widehat{H}^B_{\overline{\pi}_{h_1}}(x)$	$\frac{x+1}{r+0.5}$	$\frac{x}{r+1.5}$	$\left(\frac{\Gamma(r+1.5)x!}{\Gamma(r+1.5+q)\Gamma(x-q+1)}\right)^{\frac{1}{q}}$
$\widehat{H}^B_{\overline{\pi}_{h_2}}(x)$	$\frac{x+1}{r+3}$	$\frac{x}{r+4}$	$\left(\frac{\Gamma(r+4)x!}{\Gamma(r+4+q)\Gamma(x-q+1)}\right)^{\frac{1}{q}}$
$r^B(\Gamma_{\overline{\pi},h_1,h_2}, x)$	$\frac{2.5(x+1)}{(r+0.5)(r+3)}$	$\frac{2.5x}{(r+1.5)(r+4)}$	$\left(\frac{x!}{\Gamma(x-q+1)}\right)^{\frac{1}{q}}\left(A^{\frac{1}{q}} - B^{\frac{1}{q}}\right)$
$\widehat{H}^{PR}(x)$	$\frac{(x+1)}{2}\left(\frac{1}{r+0.5} + \frac{1}{r+3}\right)$	$\frac{x}{\sqrt{(r+4)(r+1.5)}}$	$\left(\frac{x!}{\Gamma(x-q+1)} \cdot \frac{\ln(B/A)}{1/A-1/B}\right)^{\frac{1}{q}}$

Now, applying formulas from Table 1 we can calculate Bayes and PRGM estimators under selected loss functions. Table 2 presents results.

In the above example the interesting prior and posterior distributions are easy to compute. In practice, it is not easy to compute the exact distributions and interesting posterior quantities (for example the expected value). In the next section we apply the acceptance-rejection method. The algorithm applying this method for simulation a random sample from prior and posterior distributions π_h and $\pi_h(\cdot\,|x)$, knowing distributions π and $\pi(\cdot\,|x)$, is presented in Arias-Nicolás et al. (2016).

4. The collective risk models and premium calculations, examples

Let N, $Y_1, Y_2, \dots$ be independent random variables, where N describes the number of claims and $Y_1, Y_2, \dots$ are identically distributed random variables describing severity of claims. We consider two models.

4.1. Unknown parameter θ in the Poisson model

Assume that N has the Poisson distribution with an unknown parameter $\theta > 0$ and a distribution of Y_1 is known. The parameter θ can represent a driver's propensity to make a claim and the prior indicates how that propensity is distributed throughout the population of insured drivers (see Lemaire (1979), Gómes-Déniz (2009)). Consider the premium $H(\theta)$ which is a linaer function of θ, thus $H(\theta) = t\theta$ (the net premium, the variance principle premium, the Esscher premium, the exponential premium are examples, see Boratyńska (2008)). Now, let $X_1, X_2, \dots, X_n$ be observed i.i.d. random variables with the Poisson distribution $Poiss(\theta)$ and consider following GB loss functions (for shape see Figure 1):

- the square-error loss $L_s(a,\theta) = (\theta - a)^2$,
- the LINEX loss $L_l(a,\theta) = e^{-0,5(a-\theta)} + 0,5(a-\theta) - 1$,
- the Brown loss $L_B(a,\theta) = (\ln\theta - \ln a)^2$,
- the generalized entropy losses with q equal 2, 1 and -1:

$$L_2(a,\theta) = \left(\frac{a}{\theta}\right)^2 - 2\ln\frac{a}{\theta} - 1,\; L_1(a,\theta) = \frac{a}{\theta} - \ln\frac{a}{\theta} - 1,\, L_{(-1)}(a,\theta) = \frac{\theta}{a} + \ln\frac{a}{\theta} - 1.$$

For all these loss functions it is enough to find the collective, Bayes and PRGM estimators of θ, because if $H(\theta) = t\theta$, then

$$\hat{H}_\pi^C = t\hat{\theta}_\pi^C, \qquad \hat{H}_\pi^B = t\hat{\theta}_\pi^B, \qquad \hat{H}^{PR} = t\hat{\theta}^{PR},$$

for the square-error, Brown and generalized entropy losses. For the LINEX loss $L_l(a,\theta) = e^{c(a-\theta)} + c(a-\theta) - 1$, with a constant c, we have

$$\hat{H}_\pi^C = t\hat{\theta}_{\pi,tc}^C\;, \qquad \hat{H}_\pi^B = t\hat{\theta}_{\pi,tc}^B\,, \qquad \hat{H}^{PR} = t\hat{\theta}_{tc}^{PR},$$

where $\hat{\theta}_{\pi,tc}^C$, $\hat{\theta}_{\pi,tc}^B$, $\hat{\theta}_{tc}^{PR}$ are estimators for the LINEX loss with a constant tc (see Boratyńska (2008)).

Note that the collective and the Bayes estimator of θ for loss functions L_s and $L_{(-1)}$ are equal, but PRGM estimators are different (see Table 1).

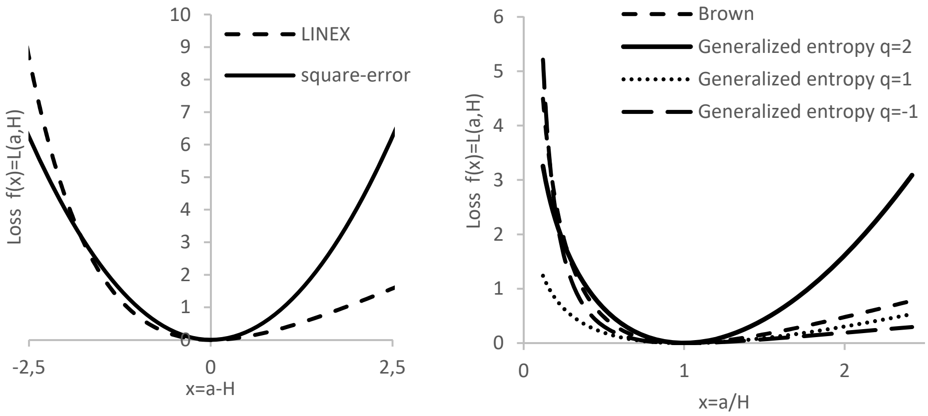

Figure 1. The graphs of loss functions

We assume that the actuary is unable to specify a simple prior distribution of the expected number of claims. Thus, let $\bar{\pi} = Gamma(3, 15)$ be the fixed prior distribution of θ with a p.d.f. $\bar{\pi}(\theta) = \frac{15^3}{2}\theta^2 \exp(-15\theta)$ for $\theta > 0$, and

$$\Gamma = \{\pi\colon\ \bar{\pi}_{h_1} \preceq_{lr} \pi \preceq_{lr} \bar{\pi}_{h_2}\}$$

be the family of priors, where

$$\bar{\pi}_{h_1}(\theta) = \frac{d}{d\theta}(1 - (1 - F_{\bar{\pi}}(\theta))^{c_1}), \qquad \bar{\pi}_{h_2}(\theta) = \frac{d}{d\theta}((F_{\bar{\pi}}(\theta))^{c_2})$$

and $c_1 = c_2 = 1.5$. Then $dK(\bar{\pi}, \bar{\pi}_{h_1}) = dK(\bar{\pi}, \bar{\pi}_{h_2}) = 0.148$. The class Γ expresses the inaccuracy in determining the cumulative distribution function of $\bar{\pi}$. The parameters c_1, c_2 provides the degree of distortion and can be elicited by fixing a reasonable distance in terms of Kolmogorov metric.

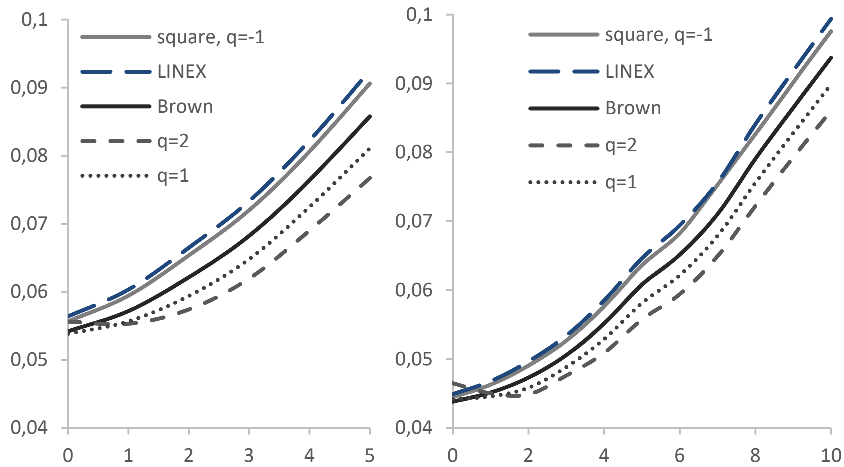

Figure 2. Oscillation $r^B(\Gamma, x)$ of Bayes estimators for different loss functions, $n = 5$ (left) and $n = 10$ (right)

Figure 2 presents the oscillation of the Bayes estimators for $n = 5$ and $n = 10$ and different values of $x = \sum_{i=1}^{n} X_i$. Table 3 shows the oscillation of the collective estimator for different losses. We see that the oscillation for Bayes estimators is an increasing function of x (except the generalized entropy loss with $q = 2$) and it is smaller than the oscillation for the collective estimators for $\frac{x}{n} < 0.5$. The greatest oscillation is for the LINEX loss.

Table 1. Oscillation of the collective estimator of θ

Loss	square	LINEX	Brown	Generalized entropy loss		
				$q = 2$	$q = 1$	$q = -1$
$r^C(\Gamma)$	0.076	0.078	0.073	0.071	0.071	0.076

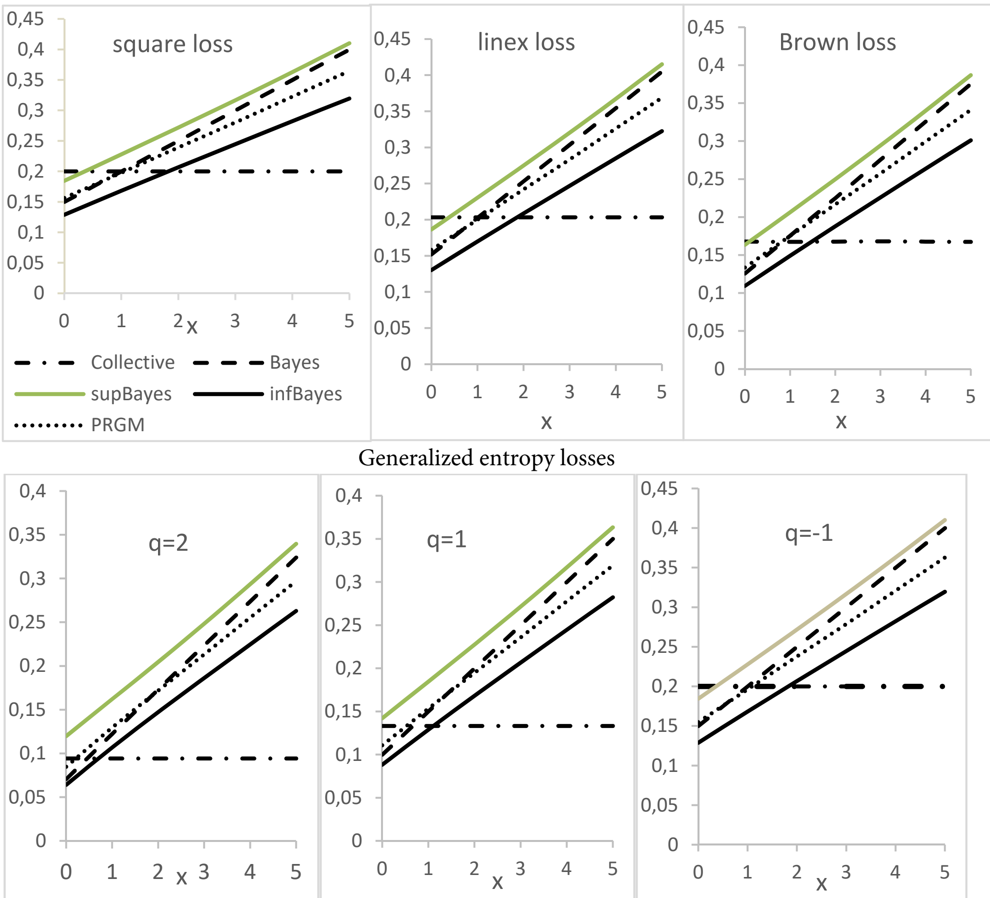

Figure 3. Values of collective, Bayes and PRGM estimators and minimum and maximum of Bayes estimators for different loss functions and $n = 5$

Figures 3 and 4 show values of minimum and maximum of Bayes estimators, collective estimators and Bayes estimators for the prior $\bar{\pi} = Gamma(3{,}15)$ and PRGM estimators for two values of n and different $x = \sum_{i=1}^{n} X_i$. The oscillation of Bayes estimators is the smallest if $\frac{x}{n}$ is closed to the expected value $E_{\bar{\pi}}\theta = 0.2$.

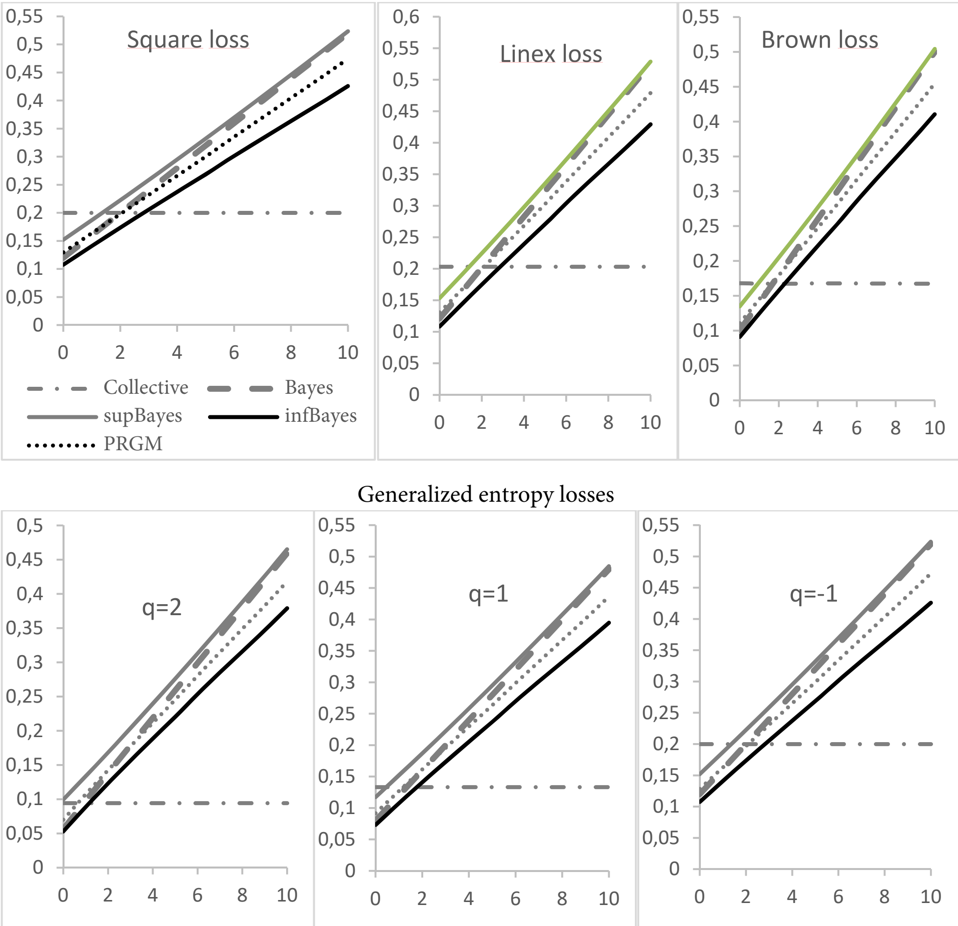

Figure 4. Values of collective, Bayes and PRGM estimators and minimum and maximum of Bayes estimators for different loss functions and $n = 10$

We use n and x small, because n is interpreted as the number of periods (years) we observe, for example, a driver, and x is the number of claims during the n periods. The prior represents the population behaviour of the parameter θ. Our results (Bayesian and PRGM premiums) have similar interpretation as the rules in the credibility theory. They combine knowledge about a single driver with knowledge about the entire population. Similar models with a parametric class of priors or an ε-contamination class of priors and the square-error loss or LINEX loss were considered in Boratyńska (2008) and Gómez-Déniz (2009).

4.2. Unknown parameters θ and λ of distributions of the number and severity of claims

Assume that random variable N has a distribution $f_1(\cdot\,|\theta)$ depending on an unknown parameter $\theta \in \Theta$, and a random variable Y_1 has a distribution $f_2(\cdot\,|\lambda)$ depending on an unknown parameter $\lambda \in \Lambda$. Consider the premium of the form

$$H(\theta,\lambda) = H_1(\theta)H_2(\lambda),$$

where H_1 and H_2 are increasing and continuous functions of θ and λ, respectively.

Let $X_1, X_2, \dots, X_n$ be observed i.i.d. random variables with a p.d.f. $f_1(\cdot\,|\theta)$ and $Z_1, Z_2, \dots, Z_m$ be observed i.i.d. random variables with a p.d.f. $f_2(\cdot\,|\lambda)$, all variables are conditionally independent, knowing parameters θ and λ. Assume that θ and λ are independent, and $\theta \sim \mu$ and $\lambda \sim \upsilon$. Denote $X = (X_1, X_2, \dots, X_n)$ and $Z = (Z_1, Z_2, \dots, Z_m)$. Let x and z be observed values of random variables X and Z. It can be seen directly that the posterior distributions $\mu(\cdot\,|x)$ and $\upsilon(\cdot\,|z)$ are independent. Consider the following GB loss functions: square-error loss, Brown loss, generalized entropy loss (see Table 1). Then, the collective and Bayes premiums are equal

$$\widehat{H}^C_{\mu,\upsilon} = \widehat{H}^C_{1,\mu}\widehat{H}^C_{2,\upsilon}, \quad \widehat{H}^B_{\mu,\upsilon}(x,z) = \widehat{H}^B_{1,\mu}(x)\widehat{H}^B_{2,\upsilon}(z).$$

Let Γ^* be a family of priors on the space $\Theta \times \Lambda$ with a p.d.f. given by

$$\pi(\lambda,\theta) = \mu(\theta)\upsilon(\lambda),$$

where

$$\mu \in \{\mu\colon \bar{\mu}_{h_1} \preceq_{lr} \mu \preceq_{lr} \bar{\mu}_{h_2}\},\ \upsilon \in \{\upsilon\colon \bar{\upsilon}_{h_3} \preceq_{lr} \upsilon \preceq_{lr} \bar{\upsilon}_{h_4}\},$$

$\bar{\mu}$ and $\bar{\upsilon}$ are fixed priors on the spaces Θ and Λ, respectively, and h_1, h_2, h_3, h_4 are fixed distortion functions (h_1, h_3 are concave and h_2, h_4 are convex). Assume that for every $\pi \in \Gamma^*$ and every x and z the Bayes premium exists. Then (applying Theorem 4) the minimum and maximum of Bayes estimators of the premium H are given by

$$\inf_{\pi\in\Gamma^*} \widehat{H}^B_\pi(x,z) = \widehat{H}^B_{1,\bar{\mu}_{h_1}}(x)\widehat{H}^B_{2,\bar{\upsilon}_{h_3}}(z), \qquad \sup_{\pi\in\Gamma^*} \widehat{H}^B_\pi(x,z) = \widehat{H}^B_{1,\bar{\mu}_{h_2}}(x)\widehat{H}^B_{2,\bar{\upsilon}_{h_4}}(z),$$

and using Theorem 2 we have the PRGM estimator of H.

Example 3. Assume that $N \sim Poiss(\theta)$ and Y_1 has an exponential distribution with a density given by $f_2(y|\lambda) = \frac{1}{\lambda}\exp(-\frac{y}{\lambda})$ for $y > 0$, depended on an unknown parameter $\lambda > 0$. Consider the net premium

$$H(\theta,\lambda) = H_1(\theta)H_2(\lambda) = \theta\lambda.$$

Assume that θ has the prior distribution $\bar{\mu} = Gamma(\alpha,\beta)$ and λ has the prior distribution $\bar{\upsilon} = IGamma(a,b)$ with a density function

$$\bar{\upsilon}(\lambda) = \frac{b^a}{\Gamma(a)}\lambda^{-a-1}\exp\left(-\frac{b}{\lambda}\right) \quad \text{for } \lambda > 0\,,$$

where α, β, a,b are fixed positive parameters and $\alpha > 2$ and $a > 1$.

If $X = x = (x_1, x_2, \dots, x_n)$ and $Z = z = (z_1, z_2, \dots, z_m)$, then the posterior distributions are $\bar{\mu}(\cdot\,|x) = Gamma(\alpha + \sum_{i=1}^{n} x_i, \beta + n)$ and $\bar{\upsilon}(\cdot\,|z) = IGamma(a + m, b + \sum_{i=1}^{m} z_i)$. We obtain the following collective and Bayes premiums:

- under the square-error loss and the generalized entropy loss for $q = -1$

$$\hat{H}^C_{\bar{\mu},\bar{\upsilon}} = \frac{\alpha b}{\beta(a-1)}, \qquad \hat{H}^B_{\bar{\mu},\bar{\upsilon}}(x,z) = \frac{(\alpha + \sum_{i=1}^{n} x_i)(b + \sum_{i=1}^{m} z_i)}{(\beta + n)(a + m - 1)},$$

- under the Brown loss

$$\hat{H}^C_{\bar{\mu},\bar{\upsilon}} = \exp(\psi(\alpha,\beta) - \psi(a,b)),$$

$$\hat{H}^B_{\bar{\mu},\bar{\upsilon}}(x,z) = \exp\left(\psi(\alpha + n, \beta + \sum_{i=1}^{n} x_i) - \psi(a + m, b + \sum_{i=1}^{m} z_i)\right),$$

where $\psi(s,t) = \int_0^{+\infty} \ln y \frac{t^s}{\Gamma(s)} y^{s-1} e^{-ty} dy$,

- under the generalized entropy loss for $q = 1$

$$\hat{H}^C_{\bar{\mu},\bar{\upsilon}} = \frac{(\alpha - 1)b}{\beta a}, \qquad \hat{H}^B_{\bar{\mu},\bar{\upsilon}}(x,z) = \frac{(\alpha + \sum_{i=1}^{n} x_i - 1)(b + \sum_{i=1}^{m} z_i)}{(\beta + n)(a + m)},$$

- under the generalized entropy loss for $q = 2$

$$\hat{H}^C_{\bar{\mu},\bar{\upsilon}} = \frac{b\sqrt{(\alpha-2)(\alpha-1)}}{\beta\sqrt{a(a+1)}}, \quad \hat{H}^B_{\bar{\mu},\bar{\upsilon}}(x,z) = \frac{\sqrt{(\alpha + \sum_{i=1}^{n} x_i - 2)(\alpha + \sum_{i=1}^{n} x_i - 1)}(b + \sum_{i=1}^{m} z_i)}{(\beta+n)\sqrt{(a+m+1)(a+m)}}.$$

For $i = 1{,}3$ and $j = 2{,}4$ define fixed numbers $c_i > 1$, $c_j > 1$ and the distortion functions

$$h_i(z) = 1 - (1-z)^{c_i}, \qquad h_j(z) = z^{c_j}.$$

Now, using the class Γ^* of priors and simmulation methods for calculation of posterior expected values (similarly as in Section 4.1), we obtain the minimum and maximum of collective and Bayes premiums and the PRGM premium.

5. Conclusions

The analysis proposed in this article was used to provide the optimal estimators of the risk premium in Bayesian models with the distorted band class of priors expressing some uncertainty in elicitation of a prior. It is an alternative to the parametric classes of priors used by practitioners. It expresses the uncertainty in determining a prior c.d.f., and that uncertainty is more realistic. The range of Bayes estimators and optimal PRGM estimators is obtained under the large family of GB loss functions, thus the practitioner can find the loss function expressing the severity of under- and over-estimation.

The numerical example presents the difference among the estimators of frequency of claims in the collective risk model under different loss functions. The last example presents the situation where we can apply results for one-dimensional parameter to the bidimensional parameter, thus we can estimate the net premium with unknown frequency and expected severity of claims. Ruggeri et al. (2021) consider the generalization of the distorted band class to the multivariate case. Applying their models we can try to describe a dependence structure between random variables θ and λ connected with frequency and severity of claims. The author believes that this topic could be expanded in the future.

References

Arias-Nicolás, J. P., Ruggeri, F., Suárez-Llorens, A., (2016). New Classes of Priors Based on Stochastic Orders and Distortion Functions, *Bayesian Analysis,* Vol. 11, pp. 1107–1136.

Balbas A., Garrido J., Mayoral S., (2009). Properties of distortion risk measures, *Methodology and Computing in Applied Probability* 11, p. 385, https://doi.org/10.1007/s11009-008-9089-z.

Berger, J. O., (1994). An overview of robust Bayesian analysis, *Test,* Vol. 3, pp. 5–124 (with discussion).

Boratyńska, A., (1997). Stability of Bayesian inference in exponential families, *Statistics & Probability Letters,* Vol. 36, pp. 173–178.

Boratyńska, A., (2002). Robust Bayesian estimation with asymmetric loss function, *Applicationes Mathematicae,* Vol. 29, pp. 297–306.

Boratyńska, A., (2008). Posterior regret Γ-minimax estimation of insurance premium in collective risk model, *ASTIN Bull.,* Vol. 38, pp. 277–291.

Boratyńska, A., (2017). Robust Bayesian estimation and prediction of reserves in exponential model with quadratic variance function, *Insurance Math. Econom.,* Vol. 76, pp. 135–140.

Chan, J., Choy, B., Makov, U., (2008). Robust Bayesian analysis of loss reserves data using the generalized t-distribution, *ASTIN Bull.,* Vol. 38, pp. 207–230.

Furman, E., Zitikis, R., (2008). Weighted premium calculation principles, *Insurance Math. Econom.***,** Vol. 42, pp. 459–465.

Gómez-Déniz, E., (2008). A generalization of the credibility theory obtained by using the weighted balanced loss function, *Insurance Math. Econom.,* Vol. 42, pp. 850–854.

Gómez-Déniz, E., (2009). Some Bayesian Credibility Premiums Obtained by Using Posterior Regret Γ-Minimax Methodology, *Bayesian Analysis,* Vol. 4, pp. 223–242.

Heilmann, W., (1989). Decision theoretic foundations of credibility theory, *Insurance Math. Econom.,* Vol. 8, pp. 77–95.

Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M., (2009). *Modern actuarial risk theory*, Springer-Verlag, Berlin Heidelberg.

Karimnezhad, A., Parsian, A., (2014). Robust Bayesian methodology with applications in credibility premium derivation and future claim size prediction, *AstA Advances in Statistical Analysis,* Vol. 98, pp. 287–303.

Karimnezhad, A., Parsian, A., (2018). Bayesian and robust Bayesian analysis in a general setting, *Commun in Statist. Theory and Methods,* Vol. 48, pp. 3899–3920.

Lemaire, J., (1979). How to define a bonus-malus system with an exponential utility function, *ASTIN Bull.,* Vol. 10, pp. 274–282.

Męczarski, M., (2015). Stochastic orders in the Bayesian framework, *Annals of the Collegium of Economic Analysis,* Vol. 37, pp. 339–360.

Ríos Insua, D., Ruggeri, F. (eds.), (2000). Robust Bayesian analysis, *Lecture Notes in Statistics,* Vol. 152, Springer-Verlag, New York.

Ríos Insua, D., Ruggeri, F., Vidakovic, B., (1995). Some results on posterior regret Γ-minimax estimation, *Statistics and Decisions,* Vol. 13, pp. 315–331.

Ruggeri, F., Sánchez-Sánchez, M., Sordo, M.A., Suárez-Llorens, A., (2021). On a New Class of Multivariate Prior Distributions: Theory and Application in Reliability, *Bayesian Analysis,* Vol. 16, pp. 31–60.

Sánchez-Sánchez, M., Sordo, M. A., Suárez-Llorens, A., Gómez-Déniz, E., (2019). Deriving robust Bayesian premiums under bands of prior distributions with applications, *ASTIN Bull.*, Vol. 49, pp. 147–168.

Shaked, M., Shanthikumar, J. G., (2007). *Stochastic Orders*, Springer, New York, USA.

Sivaganesan, S., Berger, J. O., (1989). Ranges of posterior measures for priors with unimodal contaminations, *Annals of Statist.,* Vol. 17, pp. 868–889.

Varian, H. R., (1974). A Bayesian approach to real estate assessment, *Studies in Bayesian Econometrics and Statistics,* North Holland, pp. 195–208.

Young, V., (2000). Credibility using semiparametric models and a loss function with a constancy penalty, *Insurance Math. Econom.,* Vol. 26, pp. 151–156.