On the use of posterior regret $\Gamma$-minimax actions to obtain credibility premiums

E. Gómez-Déniz$^{a,\ast}$, J.M. Pérez-Sánchez$^b$, F.J. Vázquez-Polo$^a$

$^a$Department of Quantitative Methods, University of Las Palmas de Gran Canaria, 35017-Las Palmas de Gran Canaria, Spain
$^b$Department of Quantitative Methods in Economics, University of Granada, 18011-Granada, Spain

Received January 2005; received in revised form January 2006; accepted 26 January 2006

Abstract

Computing premiums in a Bayesian context requires the use of a prior distribution that the unknown risk parameter follows in the heterogeneous portfolio. Following the methodology that an actuary only has vague information about this parameter and therefore is unable to specify a simple prior, we choose a class $\Gamma$ of priors and compute posterior regret $\Gamma$-minimax premiums which can be written, under appropriate likelihoods and priors, as a credibility formula.

JEL classification: C11

IM classification: IM30

Keywords: Bayesian methodology; Posterior regret $\Gamma$-minimax; Classes of prior distributions

1. Introduction

This article deals with the problem of estimating the Bayes premium when knowledge of the prior distribution is restricted to the fact that it is a member of a family $\Gamma$ of prior distributions. In practice, prior knowledge is often vague and any elicited prior distribution is only an approximation to the true one. Various solutions to this problem have been proposed. The literature on this subject is reviewed in Berger (1994) and Ríos and Ruggeri (2000), which are the basic references for this class of problems. One of the solutions proposed is robust Bayesian methodology, which has been applied over the last two decades and addressed recently in Actuarial Science (see Gómez et al. (2000, 2002); Ríos et al. (1999); among others).

By using robust Bayesian analysis, a range of Bayes premiums to be charged is obtained. Although the user now has a range of possible actions, perhaps it is unclear as to which of these is most appropriate. In other words, “how can we recommend a range of actions to an untrained practitioner?” The procedure we present here provides the practitioner an estimator which lies between standard and robust Bayesian methodology, and it is known as the posterior regret...
The most common procedures for determining actions among the rules in $\mathbb{P}$, the action space which is a subset of the real line $\mathbb{R}$, are the Bayes risk, minimax and $\Gamma$-minimax procedures. The Bayes criterion is used if precise information about the prior is given. On the other hand, the minimax criterion is applied if no information about the prior is available. The $\Gamma$-minimax approach (Eichenauer et al., 1988; Vidakovic, 2000) deals with partial prior information, using a class $\Gamma$ of prior distributions. In the extreme case when no information is available, the $\Gamma$-minimax setup is equivalent to the usual minimax procedure. On the other hand, if the investigator has substantial prior information, then the class may be narrow. The $\Gamma$-minimax setup becomes the usual Bayes one when the class $\Gamma$ contains a single prior.

Thus, when prior distribution is $\pi_0$, the actions $\mathcal{P} \in \mathbb{P}$ with smaller Bayes risk are preferred. The optimal procedure in this case, if it exists, is $\mathcal{P}_{\pi_0}$, which is the Bayes estimator obtained by minimizing the posterior expected loss of an action $\mathcal{P}$ under $\pi_0^x$. Here, $\pi_0^x$ is the posterior distribution of the risk parameter, $\theta \in \Theta$, given the sample information $x$.

Let $X_i, i = 1, 2, \ldots, t$ be independent and identically distributed random variables with values in $\mathcal{X} \subset \mathbb{R}$, representing the claim size of a single risk during the last $t$ periods. Let the distribution of the $X_i$ be given by $f(x \mid \theta)$ depending on $\theta$, with structure function (prior distribution) $\pi_0(\theta)$. For a given loss function $L: \Theta \times \mathbb{P} \rightarrow \mathbb{R}$, the risk function $R: \Theta \times \mathbb{P} \rightarrow \mathbb{R}$ is given by

$$R(\theta, \mathcal{P}) = \mathbb{E}_f[L(\theta, \mathcal{P})] = \int_{\mathcal{X}} L(\theta, \mathcal{P}) f(x \mid \theta) dx,$$

and the posterior expected loss of an action $\mathcal{P}$, known as the Bayes risk, is given by

$$\rho(\pi_0^x, \mathcal{P}) = \mathbb{E}_{\pi_0^x}[L(\theta, \mathcal{P})] = \int_{\Theta} L(\theta, \mathcal{P}) \pi_0^x(\theta) d\theta.$$

It is well known that in the actuarial context the unknown risk premium $\mathcal{P}$, denoted by $\mathcal{P}(\theta)$, can be obtained by minimizing the expected loss $\mathbb{E}_f[L(\theta, \mathcal{P})]$ with $\mathcal{P} \in \mathbb{P}$. If experience is not available, the actuary computes the collective premium, which is given by minimizing the risk function, i.e. minimizing $\mathbb{E}_{\pi_0}[L(\theta, \mathcal{P})]$. Finally, if experience is available, the actuary takes a sample $x$ from the random variables $X_i, i = 1, 2, \ldots, t$ and uses this information to estimate the unknown risk premium $\mathcal{P}$, obtained by minimizing the Bayes risk, i.e. minimizing $\mathbb{E}_{\pi_0^x}[L(\theta, \mathcal{P})]$.

Using the quadratic loss function, the risk, collective and Bayes premiums are given, respectively, by

$$\mathcal{P}(\theta) = \int_{\mathcal{X}} xf(x \mid \theta) dx,$$

$$\mathcal{P}_{\pi_0} = \int_{\Theta} \mathcal{P}(\theta) \pi_0(\theta) d\theta,$$

and

$$\mathcal{P}_{\pi_0^x} = \int_{\Theta} \mathcal{P}(\theta) \pi_0^x(\theta) d\theta.$$

For more details, see Gómez et al. (2000) and Heilmann (1989).

When the risk $X$ represents the claim size, the likelihood must be a continuous distribution. Examples of distributions used in these cases are normal (Heilmann, 1989; Klugman, 1992), log-normal (Hogg and Klugman, 1984; Sarabia et al., 2004), Pareto (DuMouchel and Olshen, 1974; Heilmann, 1989) and Gamma (or exponential) (Eichenauer et al., 1988; Heilmann, 1989).
In this paper we assume that the claim size follows a gamma distribution with parameters $\theta > 0$ and $\nu > 0$, $\mathcal{G}(\theta, \nu)$. Here it is assumed that $\nu$ is a fixed parameter value. Furthermore, we assume that $\theta$ follows a gamma prior distribution with parameters $a > 0$ and $b > 0$, $\mathcal{G}(a, b)$. This model has been considered by Eichenauer et al. (1988), Heilmann (1989) and Makov (1995), among others.

**Example 1.** Suppose that the claim size follows a gamma distribution with parameters $\theta > 0$ and $\nu > 0$, i.e. $f(x \mid \theta) \propto x^{\nu-1} \exp\{-\theta x\}$, while the prior distribution is again a gamma with parameters $a$ and $b$. With data $x$, the posterior distribution is again a gamma, but now with updated parameters $a + tx$ and $b + tv$.

Now, under the net premium principle and the above gamma–gamma model we have

$$P(\theta) = \frac{\nu}{\theta},$$

$$P_{\pi_0} = \frac{a}{b - 1}, \quad b > 1,$$

$$P_{\pi_0}^x = \frac{a + tx}{b + tv - 1}, \quad b + tv > 1,$$

as the risk, collective and Bayes premiums, respectively.

For a more detailed discussion of these expressions, see the papers by Eichenauer et al. (1988) and Heilmann (1989).

It is interesting to note that the Bayes premium in (2) can be rewritten as

$$P_{\pi_0}^x = \frac{a + tx}{b + tv - 1} = \frac{tv}{b + tv - 1} \bar{x} + \frac{b - 1}{b + tv - 1} \frac{av}{b - 1} = Z_t \bar{x} + (1 - Z_t) P_{\pi_0}.$$

Expression (3) is a credibility formula in the sense of the Bühlmann credibility estimate (see Herzog (1996), Chapter 8, pp. 127–150) because the Bayes premium is expressed in the form of a weighted sum of the sample data and the prior information (in this case, the collective premium) and $Z_t$ is called the credibility factor.

Using the fact that $E_f(X \mid \theta) = \frac{\nu}{\theta}$, $V_f(X \mid \theta) = \frac{\nu^2}{\theta^2}$, $E_{\pi_0}[V_f(X \mid \theta)] = v a^2 / ((b - 1)(b - 2))$ and $V_{\pi_0}[E_f(X \mid \theta)] = v^2 a^2 / ((b - 1)^2 (b - 2))$, as it is simple to prove, the credibility factor $Z_t$ can be rewritten as

$$Z_t = \frac{tv}{b + tv - 1} = \frac{t}{t + K},$$

where $K = (b - 1)/v = E_{\pi_0}[V_f(X \mid \theta)] / V_{\pi_0}[E_f(X \mid \theta)]$. Then the credibility factor is in the same form as in Heilmann (1989).

3. **Posterior regret $\Gamma$-minimax premiums**

As an alternative approach to the Bayes setup analyzed above, another procedure is that in which practitioners suppose that a correct prior $\pi_0$ exists but they are unable to apply the pure Bayesian assumption, perhaps because they are not sure of it or are unable to specify it totally and thus they assign a prior $\pi_0^*$ to the risk parameter $\theta$, which represents a well approximation of the true prior. This situation can also be considered when the question must be solved by two or more decision-makers and they do not agree about the prior distribution to be used. A common approach to prior uncertainty in Bayesian analysis is to choose a class $\Gamma$ of prior distributions and to compute the range of Bayes actions as the prior ranges over $\Gamma$. This is known as robust Bayesian methodology. An alternative consists of choosing a procedure which lies between the Bayes action and the Bayes robust methodology. This hybrid approach is known as the $\Gamma$-minimax regret principle, or the posterior regret $\Gamma$-minimax principle (PRGM, henceforth).

If $\rho(\pi_0^*, P)$ is the posterior expected loss of an action $P$ under $\pi_0^*$, the posterior regret of $P$ is defined as (Rios et al. (1995) and Zen and DasGupta (1993))

$$r(\pi_0^*, P) = \rho(\pi_0^*, P) - \rho(\pi_0^*, P_{\pi_0}),$$

which measures the loss of optimality by choosing $P$ instead of the optimal action $P_{\pi_0}$.
Now $\mathcal{P}_M \in \mathbb{P}$ is the PRGM action if
\[
\inf_{\mathcal{P} \in \mathbb{P}} \sup_{\pi_0 \in \Gamma} r(\pi^x_0, \mathcal{P}) = \sup_{\pi_0 \in \Gamma} r(\pi^x_0, \mathcal{P}_M).
\]

The PRGM procedure is based on the line that the optimal action minimizes the supremum of the function cost over distributions in class $\Gamma$. Therefore, the actuary would be wise to ensure that the largest possible increase in risk resulting from making the wrong choice of prior distribution should be kept as small as possible.

## 4. First results in the gamma–gamma model

Suppose that $f(x \mid \theta)$ is the gamma distribution and that $\theta$ is a single unknown parameter with a prior distribution $\pi_0 \in \Gamma$. In addition, it is known that $\Gamma$ is the family of gamma prior distributions (its natural conjugate), but the investigator may be unable to specify completely the parameters of this prior distribution; or perhaps two or more decision-makers may not agree as to which prior distribution should be used. Thus, we use the following classes of prior distributions:

- $\Gamma_1 = \{ \pi(\theta) = G(a, b) : a_1 \leq a \leq a_2, b \text{ fixed} \}$
- $\Gamma_2 = \{ \pi(\theta) = G(a, b) : a_1 \leq a \leq a_2, b_1 \leq b \leq b_2 \}$
- $\Gamma_3 = \{ \pi(\theta) = G(a, b) : \gamma_1 \leq \mathcal{P}_\pi \leq \gamma_2, b \text{ fixed} \}$

Class $\Gamma_3$ has the feature of dealing with moment specification; a similar class appears in Eichenauer et al. (1988). It is reasonable to consider that if the actuary has subjective knowledge about the distribution of the risk parameter, then knowledge is also available about the risk premium, a characteristic of the prior distribution.

The following result provides a guide for finding the posterior regret $\Gamma$-minimax action in the gamma–gamma model and when the net premium principle is assumed. This result is based on Proposition 2.1 in Rios et al. (1995) using the fact that the PRGM action is the midpoint of the interval $[\inf_{\pi \in \Gamma} \mathcal{P}_\pi, \sup_{\pi \in \Gamma} \mathcal{P}_\pi]$.

**Theorem 1.** Under the gamma–gamma model the PRGM net premium is given by
\[
\mathcal{P}^\Gamma_{M} = v \frac{\alpha_i + t\bar{x}}{\beta_i + tv - 1}, \quad i = 1, 2, 3,
\]
where
\[
\alpha_1 = \frac{a_1 + a_2}{2}, \quad \beta_1 = b.
\[
\alpha_2 = \frac{a_1 b_1 + a_2 b_2 + (a_1 + a_2)(tv - 1)}{b_1 + b_2 + 2(tv - 1)}, \quad \beta_2 = \frac{2b_1 b_2 + (b_1 + b_2)(tv - 1)}{b_1 + b_2 + 2(tv - 1)}.
\]
\[
\alpha_3 = \frac{(\gamma_1 + \gamma_2)(b - 1)}{2v}, \quad \beta_3 = b.
\]

**Proof.** Under quadratic loss, $r(\pi^x, \mathcal{P})$ is given by
\[
r(\pi^x, \mathcal{P}) = \rho(\pi^x, \mathcal{P}) - \rho(\pi^x, \mathcal{P}_\pi)
\]
\[
= \int_\Theta (\theta - \mathcal{P})^2 \pi^x(\theta)d\theta - \int_\Theta (\theta - \mathcal{P}_\pi)^2 \pi^x(\theta)d\theta
\]
\[
= -2\mathcal{P}\mathcal{P}_{\pi} + \mathcal{P}^2 + \mathcal{P}_{\pi}^2 = (\mathcal{P} - \mathcal{P}_{\pi})^2.
\]

Now, using the classes $\Gamma_i, i = 1, 2$, it is straightforward to obtain, using (2), that
\[
\inf_{\pi \in \Gamma_1} \mathcal{P}_{\pi} = v \frac{a_1 + t\bar{x}}{b + tv - 1}, \quad \sup_{\pi \in \Gamma_1} \mathcal{P}_{\pi} = v \frac{a_2 + t\bar{x}}{b + tv - 1}, \quad (6)
\]
\[
\inf_{\pi \in \Gamma_2} \mathcal{P}_{\pi} = v \frac{a_1 + t\bar{x}}{b_2 + tv - 1}, \quad \sup_{\pi \in \Gamma_2} \mathcal{P}_{\pi} = v \frac{a_2 + t\bar{x}}{b_1 + tv - 1}. \quad (7)
\]
For the class $\Gamma_3$, it is necessary to consider first that the moment condition, using (1), $\gamma_1 \leq \mathcal{P}_\pi \leq \gamma_2$ is equivalent to
\[
\frac{\gamma_1(b-1)}{v} \leq a \leq \frac{\gamma_2(b-1)}{v},
\]
then
\[
\inf_{\pi \in \Gamma_3} \mathcal{P}_{\pi x} = v \frac{\gamma_1(b-1)}{b + tv - 1}, \quad \sup_{\pi \in \Gamma_3} \mathcal{P}_{\pi x} = v \frac{\gamma_2(b-1)}{b + tv - 1}.
\]
(8)

Now, it is only necessary to use the fact that the posterior regret action is the midpoint of the interval $[\inf_{\pi \in \Gamma_i} \mathcal{P}_{\pi x}, \sup_{\pi \in \Gamma_i} \mathcal{P}_{\pi x}]$, $i = 1, 2, 3$, to obtain the desired result. \( \square \)

Remark 1. Notice that $\alpha_1$ is the midpoint of the interval $[\alpha_1, \alpha_2]$, while $\alpha_2 \to \frac{\alpha_1 + \alpha_2}{2}$ and $\beta_2 \to \frac{\beta_1 + \beta_2}{2}$ when $t \to \infty$.

Remark 2. It is interesting to note that the prior distributions in $\Gamma_i$, $i = 2, 3$ depend on the size of the sample, $t$, but do not depend on the sample observations. This fact is important because when the practitioner wants to estimate the parameters of the prior distribution, it will not be perturbed by the sample data.

The following corollary addresses the issue of whether PRGM premiums are Bayes.

Corollary 1. $\mathcal{P}_M^{\Gamma_i}$ are Bayes actions for $\Gamma_i$, $i = 1, 2, 3$.

Proof. This is straightforward, taking into account Proposition 3.2 in Rios et al. (1995) and the fact that closed intervals on the real line are connected sets. \( \square \)

Remark 3. It is simple to show that $\mathcal{P}_M^{\Gamma_i}, i = 1, 2, 3$, can be rewritten as a credibility formula with the credibility factor as in (4).

5. $\varepsilon$-Contaminated classes in the gamma–gamma model

It is possible to extend the above study to the $\varepsilon$-contaminated classes (Berger, 1985, 1994; Ríos and Ruggeri, 2000; Gómez et al., 2000, 2002; among others). If $\pi_0$ is the base elicited prior, the $\varepsilon$-contaminated class is given by
\[
\Gamma^\varepsilon = \{\pi : \pi = (1 - \varepsilon)\pi_0 + \varepsilon q, q \in \Gamma^*\},
\]
where $\Gamma^*$ can be a general class including the classes $\Gamma_i$, $i = 1, 2, 3$. We write $\Gamma^\varepsilon_i$ if $\Gamma^* = \Gamma_i$, $i = 1, 2, 3$.

In order to compute PRGM premiums under this class, we first need the following result.

Theorem 2. Under the $\varepsilon$-contaminated class $\Gamma^\varepsilon$, the PRGM net premium is given by
\[
\mathcal{P}_M^{\Gamma^\varepsilon} = (1 - \varepsilon)\mathcal{P}_{\pi_0^x} + \varepsilon \mathcal{P}_M^{\Gamma^*}, \quad (9)
\]
where $\mathcal{P}_{\pi_0^x}$ is the Bayes premium obtained under $\pi_0^x$, and $\mathcal{P}_M^{\Gamma^*}$ is the PRGM premium under the class $\Gamma^*$ given by
\[
\mathcal{P}_M^{\Gamma^*} = \frac{1}{2} \left( \inf_{q \in \Gamma^*} \mathcal{P}_{\pi_0^x} + \sup_{q \in \Gamma^*} \mathcal{P}_q^x \right).
\]

Proof. This is straightforward from the following result (Berger, 1985, p. 216),
\[
\inf_{\pi \in \Gamma^\varepsilon} \mathcal{P}_M^{\Gamma^\varepsilon} = (1 - \varepsilon)\mathcal{P}_{\pi_0^x} + \varepsilon \inf_{q \in \Gamma^*} \mathcal{P}_q^x,
\]
and
\[
\sup_{\pi \in \Gamma^\varepsilon} \mathcal{P}_M^{\Gamma^\varepsilon} = (1 - \varepsilon)\mathcal{P}_{\pi_0^x} + \varepsilon \sup_{q \in \Gamma^*} \mathcal{P}_q^x. \quad \square
\]

Now under the gamma–gamma model we have the following proposition.
Proposition 1. Using the classes $I_i^ε$, $i = 1, 2, 3$ and the gamma–gamma model, the PRGM net premium is given by

$$P_M^{I_i^ε} = v \left\{ (1 - ε) \frac{a + i \bar{X}}{b + tv - 1} + ε \frac{α_i + i \bar{X}}{β_i + tv - 1} \right\},$$

(10)

where $α_i, β_i, i = 1, 2, 3$, as in Theorem 1.

Proof. The proof immediately follows from Theorems 1 and 2, using expressions (5) and (9). □

Observe that this new situation can be thought of as a compromise between the pure Bayes option ($ε = 0$) and the pure PRGM criterion studied ($ε = 1$). Intermediate situations can be thought of as hybrid positions between the two procedures.

The next corollary is a direct consequence of Proposition 1 and by using $q \in I_1$.

Corollary 2. For $I_1^ε$ the PRGM net premium in Proposition 1 can be written as

$$P_M^{I_1^ε} = Z_t \bar{X} + (1 - Z_t) v \left( (1 - ε)a + εα_1 \right) \frac{b - 1}{b - 1}, \quad α_1 = (a_1 + a_2)/2,$$

(11)

i.e., is a credibility formula, with credibility factor $Z_t$ as in (4).

Proof. Using the fact that under the class $I_1^ε$ the contamination distribution is $q(θ) = G(a_1 + a_2, b)$, we have that the risk premium is given by

$$\int \frac{v}{θ} [(1 - ε)π_0(θ) + εq(θ)] dθ = (1 - ε) \frac{va}{b - 1} + ε \frac{v(a_1 + a_2)/2}{b - 1} = \frac{(1 - ε)a + ε(a_1 + a_2)/2}{b - 1} v.$$

Now, using (10) we have that

$$P_M^{I_1^ε} = \frac{vt \bar{X}}{b + tv - 1} + \frac{(1 - ε)a + ε(a_1 + a_2)/2}{b + tv - 1} v$$

$$= \frac{tv \bar{X}}{b + tv - 1} + \frac{b - 1}{b - 1} \frac{(1 - ε)a + ε(a_1 + a_2)/2}{b - 1} v$$

$$= Z_t \bar{X} + (1 - Z_t) v \left( (1 - ε)a + ε(a_1 + a_2)/2 \right) \frac{b - 1}{b - 1}. \quad □$$

6. Final remarks

In this paper we have combined the tools of standard Bayesian and global robust Bayesian analysis to obtain premiums under posterior regret gamma-minimax actions. The methodology used here is very simple and credibility formulas are straightforwardly obtained.

This technique is new in the actuarial context and presents two notable advantages. First, its application is more flexible than the $I^ε$-minimax methodology used by Eichenauer et al. (1988) to obtain a credibility formula. Second, it is a decision-making mechanism which is more suitable for a practitioner with relatively little training and who is using techniques of global robustness analysis. This last technique provides a range of premiums, and the policyholder can take a decision from among them, taking into account the PRGM procedure.

It is important to notice that this methodology is profitable in a variety of situations. In the actuarial problem discussed here, the PRGM procedure provides simple explicit rules which compare favourably with Bayes’ rules, which require more rigorous assumptions about the prior distribution.

Obviously, this technique can be extended to other pairs of priors and likelihoods. For example, Ríos et al. (1995) obtained similar results in deriving the Poisson-Gamma model, as did Zen and DasGupta (1993) with the Binomial-Beta model. Alternatively, other principles for premiums can be considered, such as the exponential or variance approaches (see Heilmann, 1989).
Acknowledgements

The authors are grateful to the referee for a very careful reading of the manuscript and suggestion which improved the paper. Research partially supported by grant from DGUI (Dirección General de Universidades e Investigación, Gobierno de Canarias, Spain) under project PI2003/033.

References