

AN ACCURATE ASYMPTOTIC APPROXIMATION FOR EXPERIENCE RATED PREMIUMS

BY

RICCARDO GATTO¹

ABSTRACT

In the Bayesian approach, the experience rated premium is the value which minimizes an expected loss with respect to a posterior distribution. The posterior distribution is conditioned on the claim experience of the risk insured, represented by a n -tuple of observations. An exact analytical calculation for the experience rated premium is possible under restrictive circumstances only, regarding the prior distribution, the likelihood function, and the loss function. In this article we provide an analytical asymptotic approximation as $n \rightarrow \infty$ for the experience rated premium. This approximation can be obtained under more general circumstances, it is simple to compute, and it inherits the good accuracy of the Laplace approximation on which it is based. In contrast with numerical methods, this approximation allows for analytical interpretations. When exact calculations are possible, some analytical comparisons confirm the good accuracy of this approximation, which can even lead to the exact experience rated premium.

KEYWORDS AND PHRASES

Class of conjugate distributions, credibility theory, Laplace approximation, loss function, posterior mode.

1. INTRODUCTION

The determination of the premium for an insured risk is one of the important problems of risk theory. In the risk process the accumulated claim process is a random sum, and premium calculation is concerned with determining the appropriate compensation process. By premium calculation we mean the determination of the adequate premium for the risk assumed by an insured individual within a collectivity. Generally, we are in position to determine the collective premium only. But if some claim experience for the given individual

¹ The author thanks an anonymous referee for thoughtful comments which improved the quality of this article.

is available, we seek to determine the premium which exploits these actual claim amounts. A premium which is based on this claim experience is called an experience rated premium or a credibility premium. Generally, the credibility theory is concerned with the question of how to use the individual claim experience in order to obtain the credibility premium. Some important pioneer contributions are Bailey (1945, 1950), Lundberg (1964), Mayerson (1964), Bichsel (1964), and Bühlmann (1964, 1967). The first Bayesian foundations for credibility were laid down by Bailey (1950), whereas Bühlmann (1967) is considered as the beginnings of sustained Bayesian research in credibility. For a recent general review about Bayesian methods in actuarial sciences, refer e.g. to Makov (2001). For textbooks, see e.g. Bühlmann (1970) and Klugman (1992).

Heilmann (1989) gives an approach to the problem of determining the experience rated premium which is based on loss functions and which allows to embed the problem into a Bayesian model. In this model, we designate by the random variable $X \geq 0$ the collective risk, which is the claim amount of the risk measured within a whole collectivity. The distribution function $F_X(x) = P[X \leq x]$ is assumed known. The collective premium is then defined as

$$\begin{aligned}
 p_C &= \operatorname{arginf}_v \int L(x, v) dF_X(x) \\
 &= \operatorname{arginf}_v E[L(X, v)],
 \end{aligned}
 \tag{1}$$

where $L : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a loss function. $L(x, v)$ is the loss incurred by a decision maker taking the action v and facing the outcome x . Here, we consider the following loss functions: the quadratic loss function $L_1(x, v) = (x - v)^2$, the exponentially scaled loss function $L_2(x, v) = (e^{\alpha x} - e^{\alpha v})^2$, $\alpha > 0$, and the exponentially tilted loss function $L_3(x, v) = e^{\alpha x}(x - v)^2$, $\alpha > 0$. These three loss functions lead respectively to the well known net, exponential and Esscher premium principles. For a graphical representation of these loss functions, see Figure 1, and for other important loss functions, refer to Heilmann (1989, Section 4). We suppose that every individual risk within a collectivity is determined by an unknown parameter θ which is the realization of a random variable Θ , with known structure or prior distribution $F_\Theta(\theta) = P[\Theta \leq \theta]$. With this, we define the individual claim amount distribution as $F_{X|\Theta}(x|\theta) = P[X \leq x | \Theta = \theta]$. In analogy with the collective premium (1), we define the individual premium as

$$\begin{aligned}
 p_I(\theta) &= \operatorname{arginf}_v \int L(x, v) dF_{X|\Theta}(x|\theta) \\
 &= \operatorname{arginf}_v E[L(X, v) | \Theta = \theta].
 \end{aligned}
 \tag{2}$$

Consider the random variables X, X_1, \dots, X_n defined so that, given $\Theta = \theta$, X, X_1, \dots, X_n are independent with distribution $F_{X|\Theta}(x|\theta)$. An individual claim experience based on the last n periods is then $(X_1, \dots, X_n) | \Theta = \theta$. Hence, for a fixed individual, the claim amounts of consecutive periods are independent and identically distributed. Let $\underline{y} = (y_1, \dots, y_n)$ denote an individual observed claim experience during the last n periods, i.e. a realization of the conditional random vector $(X_1, \dots, X_n) | \Theta = \theta$. The posterior distribution of Θ for this individual with experience \underline{y} is given by $F_{\Theta|\underline{X}}(\theta|\underline{y}) = P[\Theta \leq \theta | \underline{X} = \underline{y}]$. For this

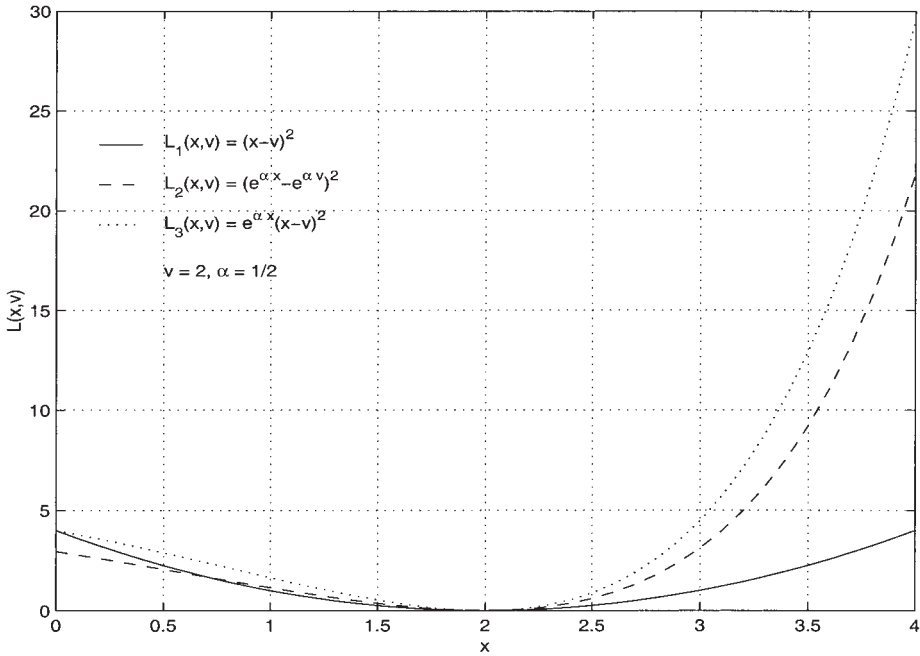


FIGURE 1: Loss functions $L_1(x, v) = (x - v)^2$, $L_2(x, v) = (e^{\alpha x} - e^{\alpha v})^2$, and $L_3(x, v) = e^{\alpha x}(x - v)^2$.

same individual, the experience rated premium or credibility premium is defined by

$$\begin{aligned}
 p_E(\underline{y}) &= \operatorname{arginf}_v \int L(p_I(\theta), v) dF_{\Theta|X}(\theta | \underline{y}) \\
 &= \operatorname{arginf}_v E[L(p_I(\Theta), v) | \underline{X} = \underline{y}].
 \end{aligned}
 \tag{3}$$

Note that in (3), the loss function which appears directly is not necessarily the same as the one which appears indirectly, through the function $p_I(\theta)$. Viewed as a function of y , $p_E(\underline{y})$ is a Bayes rule for estimating $p_I(\theta)$, see e.g. Rohatgi (1976, Section 8.8). For more details about this summary, refer to Heilmann (1989).

In order to compute the experience rated premium, we need to evaluate the integral in (3). Unfortunately, this can be solved analytically under rather restrictive circumstances only, which concern the choice of the loss function, and the choices of the prior F_{Θ} and of the likelihood $F_{X|\Theta}$, within classes of conjugate distributions. For the definition of a conjugate class and for some examples, see e.g. Heilmann (1989, Section 3). Note that under these circumstances, the experience rated premium can sometimes be re-expressed as a credibility formula, namely in the form of $p_E(\underline{y}) = z_n p_I(g(\bar{y})) + (1 - z_n)E[X]$, where $z_n \in (0, 1)$ is called credibility factor, satisfying $z_n \leq z_{n+1}, \forall n$, where $\bar{y} = n^{-1} \sum_{i=1}^n y_i$, and where $g(\bar{y})$ is a maximum likelihood estimate of θ . However, the choices of the prior, of the likelihood and of the loss function should be based

on objective criteria, possibly related with the real environment of the insurer, rather than on the feasibility of the ensuing calculations. Our scope is to provide a method which is valid under more general circumstances. Of course, numerical integration or simulation provide the most general solutions, but they do not yield an analytical result which would allow for interpretations. In this context, the Laplace method gives an asymptotic approximation of the integral appearing in (3) with a small asymptotic relative error, and this allows to obtain an accurate asymptotic approximation to the experience rated premium. The Laplace approximation has already proved to be an useful method in Bayesian statistics, see Tierney and Kadane (1986).

The rest of this article is organized as follows. The application of the Laplace approximation to the experience rated premium is described in Section 2. In Section 3, we compute analytically some approximated experience rated premiums with the technique introduced in Section 2. We mainly consider cases where the exact experience rated premium can be also obtained, in order to make analytical comparisons. These comparisons show that our approximated experience rated premiums are close to the exact ones. We also give an example where the approximation leads to the exact experience rated premium.

2. THE ASYMPTOTIC APPROXIMATION

The following theorem is based on the Laplace method of asymptotic analysis and it gives an accurate approximation for the experience rated premium. Two alternative approximations are given by Remarks 3 and 4 following Theorem 1 below.

Theorem 1. Let y denote a realization of X given a fixed value of Θ , and suppose that $F_{\Theta|X}(\theta|y)$ is an absolutely continuous distribution with density $f_{\Theta|X}(\theta|y)$ which is > 0 over the interval (a,b) . If $f_{\Theta|X}(\theta|y)$ is differentiable over (a,b) , if $f_{\Theta|X}(\theta|y)$ has a unique maximum at the interior point $\theta = \tilde{\theta}$ of (a,b) , if $\tilde{\theta}$ is the only point in (a,b) satisfying $(\partial/\partial\theta)f_{\Theta|X}(\theta|y) = 0$, if $(\partial/\partial\theta)^2 f_{\Theta|X}(\tilde{\theta}|y)$ exists and is < 0 , and if we define

$$p_L(y) = \operatorname{arginf}_v L(p_I(\tilde{\theta}),v),$$

then, $\forall y \in R_+^n$,

$$p_L(y) \sim p_E(y), \text{ as } n \rightarrow \infty, \tag{4}$$

where $p_I(\theta)$ is given by (2) and $p_E(y)$ by (3).

Note that $\tilde{\theta}$ is the posterior mode estimator (PME) of Θ , given $X = y$.

Proof. Denote by

$$S_n(v) = \int L(p_I(\theta),v) dF_{\Theta|X}(\theta|y),$$

by $g(\theta) = -n^{-1} \log f_{\Theta|X}(\theta|y)$, and by $h(\theta, v) = L(p_I(\theta), v)$. We can thus write

$$S_n(v) = \int h(\theta, v) e^{-ng(\theta)} d\theta, \quad v \geq 0. \tag{5}$$

If $f_{\Theta|X}(\theta|y)$ satisfies the hypotheses of this theorem, then $g(\theta)$ satisfies the following hypotheses: $g(\theta)$ is differentiable over (a, b) , $g(\theta)$ has a unique minimum at the interior point $\theta = \tilde{\theta}$ of (a, b) , $\tilde{\theta}$ is the only point satisfying $(\partial/\partial\theta)g(\theta) = 0$, $(\partial/\partial\theta)^2 g(\tilde{\theta}) = -n^{-1}(\partial/\partial\theta)^2 f_{\Theta|X}(\tilde{\theta}|y) / f_{\Theta|X}(\tilde{\theta}|y)$ exists and is > 0 . These hypotheses are required in order to apply the Laplace method to $S_n(v)$, which leads to the Laplace approximation

$$S_n(v) = \left(\frac{\pi}{2ng''(\tilde{\theta})} \right)^{\frac{1}{2}} h(\tilde{\theta}, v) e^{-ng(\tilde{\theta})} + O\left(\frac{e^{-ng(\tilde{\theta})}}{n^{\frac{3}{2}}} \right), \quad \text{as } n \rightarrow \infty, \tag{6}$$

where $g''(\theta) = (\partial/\partial\theta)^2 g(\theta)$. The Laplace approximation is established by an appropriate change of variables and a series expansion, both applied to the integrand in (5), followed by an application of Watson Lemma; see e.g. Bleistein and Handelsman (1986, Section 5.1). From (6) it follows that

$$\begin{aligned} \operatorname{arginf}_v S_n(v) &= \operatorname{arginf}_v [h(\tilde{\theta}, v) \{1 + O(n^{-1})\}] \\ &= \operatorname{arginf}_v [L(p_I(\tilde{\theta}), v) \{1 + O(n^{-1})\}] \\ &\sim \operatorname{arginf}_v L(p_I(\tilde{\theta}), v), \quad \text{as } n \rightarrow \infty. \end{aligned} \tag{7}$$

Note that in general the error term in (6) involves v , otherwise the asymptotic approximation above would be an exact one.

The asymptotic equivalence (7) can be detailed as follows. Denote by $\tilde{S}_n(v)$ the Laplace approximation to $S_n(v)$, i.e. the first term in the right-hand-side of (6), $v_n = \operatorname{arginf}_v S_n(v)$ and $\tilde{v}_n = \operatorname{arginf}_v \tilde{S}_n(v)$. We see that $\tilde{S}_n^{(j)}(v) = O(n^{-1/2} e^{-ng(\tilde{\theta})})$, $j = 0, 1, \dots$, where $f^{(j)}$ denotes the j -th derivative of a function f , and we can also deduce that $S_n^{(j)} = O(n^{-1/2} e^{-ng(\tilde{\theta})})$, $j = 0, 1, \dots$. From Taylor expansions, we have

$$S_n(\tilde{v}_n) - S_n(v_n) = \frac{1}{2}(\tilde{v}_n - v_n)^2 S_n''(v_n) + o([\tilde{v}_n - v_n]^2 n^{-1/2} e^{-ng(\tilde{\theta})}),$$

and

$$\tilde{S}_n(v_n) - \tilde{S}_n(\tilde{v}_n) = \frac{1}{2}(\tilde{v}_n - v_n)^2 \tilde{S}_n''(\tilde{v}_n) + o([\tilde{v}_n - v_n]^2 n^{-1/2} e^{-ng(\tilde{\theta})}).$$

By summing both equations we obtain

$$(\tilde{v}_n - v_n)^2 \frac{1}{2} [S_n''(v_n) + \tilde{S}_n''(\tilde{v}_n)] + o([\tilde{v}_n - v_n]^2 n^{-1/2} e^{-ng(\tilde{\theta})}) = O(n^{-\frac{3}{2}} e^{-ng(\tilde{\theta})}),$$

which implies

$$\tilde{v}_n = v_n + O(n^{-\frac{1}{2}}), \quad \text{as } n \rightarrow \infty.$$

This justifies (7) and completes the proof. □

Remark 1. An important advantage of the approximation $p_L(y)$ for the experience rated premium, with respect to e.g. numerical integration or simulation, is that it leads to accurate numerical results often through a simple closed form expression, which is very useful for deriving other related properties or for interpretations. In the examples presented in Section 3, one can for instance directly see the effect of the prior parameters on the experience rated premium.

Remark 2. There are other ways of expressing the integral $S_n(v)$ in the form suggested by (5). The function g is the central element which determines the quality of the Laplace approximation by means of its critical point $\hat{\theta}$. The derivation of the Laplace approximation is based on expansions of both g and h around $\hat{\theta}$. Two helpful guidelines for an appropriate identification of g and h are the following: h must be dominated by e^{-ng} , and h should be a slowly varying function.

A consequence of this is that the Laplace approximation and its ensuing premium $p_L(y)$ may not always be accurate when using the exponentially scaled loss function $L_2(x, v) = (e^{\alpha x} - e^{\alpha v})^2$ and the exponentially tilted loss function $L_3(x, v) = e^{\alpha x}(x - v)^2$ with large values of α ; for an illustration of this refer to Example 5 in Section 3.

Remark 3. An alternative approximation is as follows. Define

$$S_n^\dagger(v) = \int L(p_I(\theta), v) f_\Theta(\theta) e^{\lambda(\theta|y)} d\theta,$$

where f_Θ is the prior density of Θ , and $\lambda(\theta|y) = \sum_{i=1}^n \log f_{X_i|\Theta}(y_i|\theta)$ is the logarithmic likelihood. Clearly, $S_n^\dagger(v)$ is equal to $S_n(v)$, up to a factor independent of v . If we denote by $g^\dagger(\theta) = -n^{-1}\lambda(\theta|y)$, and by $h^\dagger(\theta, v) = L(p_I(\theta), v) f_\Theta(\theta)$, we retrieve the suitable expression for applying the Laplace method

$$S_n^\dagger(v) = \int h^\dagger(\theta, v) e^{-ng^\dagger(\theta)} d\theta. \tag{8}$$

Omitting details, we are finally led to the approximation

$$p_E(y) \sim \operatorname{arginf}_v L(p_I(\hat{\theta}), v), \text{ as } n \rightarrow \infty,$$

where $\hat{\theta}$ is the solution of the likelihood equation $(\partial/\partial\theta)\lambda(\theta|y) = 0$, therefore the maximum likelihood estimator. The hypotheses required on g^\dagger follow from some standard regularity conditions on the likelihood, namely: $\lambda(\theta|y)$ is concave in θ , the maximum likelihood estimator is the solution of the likelihood equation, and the observed Fisher information $-(\partial/\partial\theta)^2\lambda(\hat{\theta}|y)$ exists. The error of this alternative approximation is of the same order as the one of the approximation of Theorem 1, but the form of the approximation is not satisfactory, because all the prior information is kept away from the exponential appearing in (8) and this information is finally lost.

Remark 4. Another approximation can be obtained by defining $g^*(\theta, v) = -n^{-1}[\lambda(\theta|y) + \log f_\Theta(\theta) + L(p_I(\theta), v)]$, which leads to the expression

$$S_n(v) = \int e^{-ng^*(\theta,v)} d\theta,$$

also suitable for the Laplace method. If g^* satisfies the conditions expressed in terms of the function g in the proof of Theorem 1, we obtain

$$p_E(y) \sim \operatorname{arginf}_v [\{g^{*''}(\hat{\theta}(v),v)\}^{-\frac{1}{2}} e^{-ng^*(\hat{\theta}(v),v)}], \text{ as } n \rightarrow \infty, \tag{9}$$

where $\hat{\theta}(v)$ is solution of $g^{*'}(\hat{\theta}(v),v) = 0$, and where $g^{*'}(\theta,v) = (\partial/\partial\theta)g^*(\theta,v)$ and $g^{*''}(\theta,v) = (\partial/\partial\theta)^2 g^*(\theta,v)$. The approximation (9) to $p_E(y)$ depends on the prior information, and it is based on the slow varying function $h = 1$. This should give accurate approximations, but the drawback comes from the complexity of the solution, which does in general not admit a closed form, consequently making interpretations difficult.

Remark 5. Posterior densities which are unimodal are often approximated by a normal density centered at the posterior mode. Indeed, as $\theta \rightarrow \tilde{\theta}$,

$$\begin{aligned} \log f_{\Theta|X}(\theta|y) &= \log f_{\Theta|X}(\tilde{\theta}|y) \\ &+ \frac{1}{2}(\theta - \tilde{\theta})^2 \left(\frac{d}{d\theta}\right)^2 \log f_{\Theta|X}(\theta|y) \Big|_{\theta=\tilde{\theta}} + o\left([\theta - \tilde{\theta}]^2\right), \end{aligned}$$

which directly implies

$$f_{\Theta|X}(\theta|y) \sim \frac{1}{\sqrt{2\pi\sigma^2(y)}} e^{-\frac{1}{2} \frac{(\theta - \tilde{\theta})^2}{\sigma^2(y)}}, \text{ as } n \rightarrow \infty,$$

where $\sigma^2(y) = -[(d/d\theta)^2 \log f_{\Theta|X}(\theta|y)|_{\theta=\tilde{\theta}}]^{-1}$. The Laplace approximation of Theorem 1 roughly consists on approximating g by a quadratic function around its minimum and on integrating over a small neighborhood of this minimum only. The expansion above shows that this is equivalent to integrating with respect to a normal distribution centered at its mode, only locally around this mode. The small asymptotic error of the Laplace approximation results from the fact that the normal approximation is considered only locally around its center, where it is the most reliable.

The normal approximation above illustrates also that the expectation appearing in $p_E(y)$ is taken with respect to a distribution which tends to a point mass in $\tilde{\theta}$, and because for the loss functions considered $\operatorname{arginf}_v L(x,v) = x$, the approximation $p_I(\tilde{\theta})$ to $p_E(y)$ can be admitted quite intuitively. But the proof via the Laplace approximation provides a formal justification, and it also gives the asymptotic error $O(n^{-1/2})$ for the experience rated premium.

3. APPLICATIONS

In this section we apply the approximation to the experience rated premium given by Theorem 1 in various examples with different classes of conjugate

distributions, and different loss functions. In Example 1, our approximation leads to the exact experience rated premium. In Example 2 to 5, the approximated premium rates admit closed forms which allow for direct interpretations. Some small numerical examples suggest that the relative error of this approximation is around 5% when the claim experience period is around $n = 20$. Example 6 represents a situation where no simple form for the exact premium is available, whereas our approximations yields again a simple and practical formula.

Example 1. *Gamma-exponential conjugate class, quadratic loss*

Suppose that the prior distribution of Θ is $\text{Gamma}(a, b)$, with density $b^a \Gamma^{-1}(a) e^{-b\theta} \theta^{a-1}$, $\theta > 0$, and that the likelihood is $\text{Exponential}(\theta)$, having density $\theta e^{-\theta x}$, $x > 0$. It follows that the posterior distribution of $\Theta | \underline{X} = \underline{y}$ is $\text{Gamma}(a + n, b + n\bar{y})$. Consider also the quadratic loss function $L_1(x, v) = (x - v)^2$. With this choice, it is straightforward to obtain the individual premium

$$p_I(\theta) = E[X | \Theta = \theta] = \frac{1}{\theta}.$$

The exact experience rated premium is

$$p_E(\underline{y}) = E[p_I(\Theta) | \underline{X} = \underline{y}] = \frac{b + n\bar{y}}{a + n - 1}.$$

As the mode of $\text{Gamma}(\alpha, \beta)$ is $(\alpha - 1)/\beta$, it follows that the posterior distribution of $\Theta | \underline{X} = \underline{y}$ satisfies the conditions of Theorem 1 and that, given $\underline{X} = \underline{y}$, the PME of Θ is

$$\tilde{\theta} = \frac{a + n - 1}{b + n\bar{y}}.$$

With this, we easily obtain the approximation to the exact premium

$$p_L(\underline{y}) = p_I(\tilde{\theta}) = \frac{b + n\bar{y}}{a + n - 1}.$$

Although Theorem 1 tells us that $p_E(\underline{y}) \sim p_L(\underline{y})$, we have here $p_E(\underline{y}) = p_L(\underline{y})$, namely that our approximation yields the exact experience rated premium. In this case we obtain also the simple credibility formula $p_E(\underline{y}) = p_L(\underline{y}) = z_n \bar{y} + (1 - z_n) E[X]$, where $z_n = n/(a + n - 1)$ and $E[X] = b/(a - 1)$.

Example 2. *Gamma-Poisson conjugate class, quadratic loss*

Suppose that the prior distribution of Θ is $\text{Gamma}(a, b)$ and that the likelihood is $\text{Poisson}(\theta)$, with probability function $\theta^x (x!)^{-1} e^{-\theta}$, $x = 0, 1, \dots$. It follows that the posterior distribution of $\Theta | \underline{X} = \underline{y}$ is $\text{Gamma}(a + n\bar{y}, b + n)$. We also consider the quadratic loss function $L_1(\bar{x}, v) = (x - v)^2$. Simple algebra leads to

$$p_I(\theta) = E[X | \Theta = \theta] = \theta,$$

and to

$$p_E(\underline{y}) = E[p_I(\Theta) | \underline{X} = \underline{y}] = \frac{a + n\bar{y}}{b + n}.$$

The posterior distribution satisfies the conditions of Theorem 1 whenever $a + n\bar{y} > 1$. The PME of Θ is

$$\tilde{\theta} = \frac{a + n\bar{y} - 1}{b + n}. \tag{10}$$

Hence, our approximation is

$$p_L(\underline{y}) = p_I(\tilde{\theta}) = \frac{a + n\bar{y} - 1}{b + n},$$

which is close to the exact solution for a sufficiently large value of n . For a numerical illustration, we choose $a = b = 1$ and $\bar{y} = 1$. It then follows that $p_E(\underline{y}) = 1$ and that $p_L(\underline{y}) = n / (n + 1)$. The relative error of our approximation, defined as $RE_n = |p_L(\underline{y}) - p_E(\underline{y})| / p_E(\underline{y})$, is $RE_n = | -1 / (n + 1) |$. For example, $RE_9 = 10\%$ and $RE_{19} = 5\%$.

Example 3. *Beta-binomial conjugate class, quadratic loss*

Suppose that the prior distribution of Θ is $\text{Beta}(a, b)$, having density $B^{-1}(a, b) \theta^{a-1} (1 - \theta)^{b-1}$, $0 < \theta < 1$, where $B(a, b) = \int_0^1 \theta^{a-1} (1 - \theta)^{b-1} d\theta$ is the beta function, and that the likelihood is $\text{Binomial}(m, \theta)$, having probability function $m! / [(m - x)! x!]^{-1} \theta^x (1 - \theta)^{m-x}$, $x = 0, \dots, m$. It follows that the posterior distribution of $\Theta | \underline{X} = \underline{y}$ is $\text{Beta}(a + n\bar{y}, b + mn - n\bar{y})$. Consider also that the quadratic loss function $L_1(x, v) = (x - v)^2$. Simple algebra leads to

$$p_I(\theta) = E[X | \Theta = \theta] = m\theta,$$

and to

$$p_E(\underline{y}) = E[p_I(\Theta) | \underline{X} = \underline{y}] = m \frac{a + n\bar{y}}{a + b + mn}.$$

From the fact that the mode of $\text{Beta}(\alpha, \beta)$ is $(\alpha - 1) / (\alpha + \beta - 2)$, it follows that the posterior distribution of $\Theta | \underline{X} = \underline{y}$ satisfies the conditions of Theorem 1 whenever $a + n\bar{y} > 1$ and $b + mn - n\bar{y} > 1$. The PME of Θ is

$$\tilde{\theta} = \frac{a + n\bar{y} - 1}{a + b + mn - 2}. \tag{11}$$

Our approximation to the exact premium is then

$$p_L(\underline{y}) = p_I(\tilde{\theta}) = m \frac{a + n\bar{y} - 1}{a + b + mn - 2},$$

which is close to $p_E(\underline{y})$ for sufficiently large values of n . As an example, we choose $a = b = m = 2$ and $\bar{y} = 1$. It then follows that $p_E(\underline{y}) = 1$ and that $p_L(\underline{y}) = 1$, so that the relative error is zero.

Example 4. *Beta-negative binomial conjugate class, quadratic loss*

Suppose that the prior distribution of Θ is Beta(a, b) and that the likelihood is negative binomial with parameters r and θ , having probability function $(r + x - 1)! [x!(r - 1)!]^{-1} \theta^r (1 - \theta)^x$, $x = 0, 1, \dots$. It follows that the posterior distribution of $\Theta | \underline{X} = \underline{y}$ is Beta($a + nr, b + n\bar{y}$). Consider also that the quadratic loss function $L_1(x, \underline{y}) = (x - \underline{y})^2$. Then,

$$p_I(\theta) = E[X | \Theta = \theta] = r \frac{1 - \theta}{\theta},$$

and to

$$p_E(\underline{y}) = E[p_I(\Theta) | \underline{X} = \underline{y}] = r \frac{b + n\bar{y}}{a + nr - 1}.$$

The conditions Theorem 1 are satisfied whenever $b + n\bar{y} > 1$, and the PME of Θ is

$$\tilde{\theta} = \frac{a + nr - 1}{a + nr + b + n\bar{y} - 2}.$$

Our approximation is

$$p_L(\underline{y}) = p_I(\tilde{\theta}) = r \frac{b + n\bar{y} - 1}{a + nr - 1},$$

which is close to the exact solution $p_E(\underline{y})$ for a sufficiently large value of n . As a numerical example, let us choose $a = b = 2$ and $\bar{y} = 1$. It then follows that $p_E(\underline{y}) = (n + 2) / (n + 1)$ and $p_L(\underline{y}) = 1$, and that the relative error of our approximation is $RE_n = |-1 / (n + 2)|$. For example, $RE_8 = 10\%$ and $RE_{18} = 5\%$.

A consequence of both the exponentially scaled and exponentially tilted loss functions is to penalize a too low premium more than a too high premium, whereas the quadratic loss function attributes the same loss to premiums which are too high or too low. This is shown in Figure 1: with the choices $\alpha = 1/2$ and $\nu = 2$, $L_2(x, \underline{y})$ and $L_3(x, \underline{y})$ increase much faster than $L_1(x, \underline{y})$ on the right part of the graphic, but behave similarly on the left part.

Example 5. *Gamma-Poisson conjugate class, exponentially scaled loss*

Suppose that the prior distribution of Θ is Gamma(a, b) and that the likelihood is Poisson(θ). It follows that the posterior distribution of $\Theta | \underline{X} = \underline{y}$ is Gamma($a + n\bar{y}, b + n$). Consider now the exponentially scaled loss function $L_2(x, \underline{y}) = (e^{\alpha x} - e^{\alpha \underline{y}})^2$. Simple algebra leads to

$$p_I(\theta) = \frac{1}{\alpha} \log E[e^{\alpha X} | \Theta = \theta] = \frac{\theta}{\alpha} (e^\alpha - 1),$$

and to

$$p_E(\underline{y}) = \frac{1}{\alpha} (a + n\bar{y}) \log \frac{b + n}{b + n - e^\alpha + 1}.$$

The conditions of Theorem 1 are fulfilled if $a + n\bar{y} > 1$, and the PME of Θ is given by (10) in Example 2. Our approximation leads to

$$p_L(\underline{y}) = p_I(\tilde{\theta}) = \frac{1}{\alpha} \frac{a + n\bar{y} - 1}{b + n} (e^\alpha - 1).$$

It is easy to show that

$$\log \frac{b + n}{b + n - e^\alpha + 1} = \frac{e^\alpha - 1}{b + n} + O(n^{-2}), \text{ as } n \rightarrow \infty,$$

which confirms the asymptotic equivalence between the exact experience rated premium $p_E(\underline{y})$ and our approximation $p_L(\underline{y})$. In the numerical illustration, let us choose $a = \bar{b} = 1$, $\bar{y} = 1$, and $\alpha = 1/2$; this latter choice corresponds to the loss function $L_2(x, v)$ of Figure 1. It follows that $p_E(\underline{y}) = 2(n + 1) \log \{(n + 1) / (n + 2 - e^{1/2})\}$ and $p_L(\underline{y}) = 2n / (n + 1) (e^{1/2} - 1)$. Two values of the relative errors are $RE_{12} = 10\%$ and $RE_{25} = 5\%$. Note that the latter asymptotic equivalence holds also as $\alpha \rightarrow 0$, with n fixed, or with both $\alpha \rightarrow 0$ and $n \rightarrow \infty$. The accuracy of $p_L(\underline{y})$ improves with smaller α , because $L_3(x, v)$ becomes a slowly varying function; refer to Remark 2 in Section 2. In practice, small values of α are sufficient to generate asymmetric loss functions, as shown by Figure 1.

Example 6. *Beta-binomial conjugate class, exponentially tilted loss*

Suppose that the prior distribution of Θ is $\text{Beta}(a, b)$, and that the likelihood is $\text{Binomial}(m, \theta)$. It follows that the posterior distribution of $\Theta | \underline{X} = \underline{y}$ is $\text{Beta}(a + n\bar{y}, b + mn - n\bar{y})$. Consider here the exponentially tilted loss function $L_3(x, v) = e^{\alpha x}(x - v)^2$, $\alpha > 0$. With this choice, we easily find the individual premium

$$p_I(\theta) = \frac{E[Xe^{\alpha X} | \Theta = \theta]}{E[e^{\alpha X} | \Theta = \theta]} = \frac{m\theta e^\alpha}{1 + \theta(e^\alpha - 1)}.$$

The exact experience rated premium is

$$p_E(\underline{y}) = \frac{E[p_I(\Theta)e^{\text{opt}_I(\Theta)} | \underline{X} = \underline{y}]}{E[e^{\text{opt}_I(\Theta)} | \underline{X} = \underline{y}]},$$

but the integrals of both expectations above cannot be analytically evaluated. The PME of Θ was already given by (11) in Example 3, and the conditions of Theorem 1 are satisfied. With this, our approximation to the exact premium leads to

$$p_L(\underline{y}) = p_I(\tilde{\theta}) = \frac{me^\alpha(a + n\bar{y} - 1)}{a + b + mn - 2 + (e^\alpha - 1)(a + n\bar{y} - 1)}.$$

In this example, we cannot provide an analytical comparison because the exact experience rated premium cannot be computed, whereas our approximation gives the above formula.

REFERENCES

- BAILEY, A.L. (1945) "A generalized theory of credibility", *Proceedings of the Casualty Actuarial Society*, **32**, 13-20.
- BAILEY, A.L. (1950) "Credibility procedures, Laplace's generalizations of Bayes rule and the combination of collateral knowledge with observed data", *Proceedings of the Casualty Actuarial Society*, **37**, 7-23.
- BICHSEL, F. (1964) "Erfahrungstarifizierung in der Motorfahrzeughaftpflicht-Versicherung", *MSV*, **64**, 119-130.
- BLEISTEIN, N. and HANDELSMAN, R.A. (1986) *Asymptotic Expansions of Integrals*, second edition, Dover.
- BÜHLMANN, H. (1964) "Optimale Prämienstufensysteme", *MSV*, **64**, 193-214.
- BÜHLMANN, H. (1967) "Experience rating and credibility", *Astin Bulletin*, **4**, 199-207.
- BÜHLMANN, H. (1970) *Mathematical Methods in Risk Theory*, Springer.
- HEILMANN, W.-R. (1989) "Decision theoretic foundations of credibility theory", *Insurance: Mathematics and Economics*, **8**, 77-95.
- KLUGMAN, S.A. (1992) *Bayesian Statistics in Actuarial Sciences*, Kluwer.
- LUNDBERG, O. (1964), *On random processes and their application to sickness and accidents statistics*, Ph.D thesis, University of Stockholm (1940), second edition, Almquist & Wiksells.
- MAKOV, U.E. (2001) "Principal applications of Bayesian methods in actuarial science: a perspective", with discussions, *North American Actuarial Journal*, **5**, 53-73.
- MAYERSON, A.L. (1964) "The uses of credibility in property insurance ratemaking", *Giornale dell'Istituto Italiano degli Attuari*, **27**, 197-218.
- ROHATGI, V.K. (1976), *An Introduction to Probability Theory and Mathematical Statistics*, Wiley.
- TIERNEY, L. and KADANE, J.B. (1986), "Accurate approximations for posterior moments and marginal densities", *Journal of the American Statistical Association*, **81**, 82-86.

RICCARDO GATTO

Department of Mathematical Statistics and Actuarial Science,

University of Bern,

Sidlerstrasse 5, 3012 Bern

Switzerland.

Internet: gatto@stat.unibe.ch;

http://www.stat.unibe.ch/~gatto