

# Experience Rating in General Insurance by Credibility Estimation\*

Xian Zhou

Department of Applied Finance and Actuarial Studies  
Macquarie University, Sydney, Australia

## Abstract

This work presents a new credibility estimation of the probability distributions of risks under Bayes settings in a nonparametric framework without any assumption on the prior distribution. The method is then applied to general insurance premium pricing, a procedure commonly known as *experience rating*, which utilizes the insured's past claim experience to calculate the premium under a given premium principle (also referred to as a *risk measure*). By estimating the entire probability distributions of losses, our method provides a unified nonparametric framework to experience rating under arbitrary premium principles. This encompasses the advantages of the well-known approaches introduced by Bühlmann and Ferguson, while overcomes their drawbacks.

---

\*This is a joint work with Prof. Xiaoqiang Cai at the Chinese University of Hong Kong, Dr. Limin Wen at Jiangxi Normal University, and Prof Xianyi Wu at East China Normal University.

# 1 Introduction

- Pricing or measuring risks is one of the central tasks of financial enterprises and regulators in risk management.
- Numerous risk measures have been proposed for this purpose, such as value-at-risk, conditional value-at-risk, coherent risk measures, and distortion risk measures.
- In compensating insured losses in property, business, employment, health, etc., pricing risks is carried out by *premium principle*, which is considered as the translation of risk measures into the general insurance context.
- In this talk, the terms “risk measure” and “premium principle” are alternately used to indicate a rule  $H(X)$  that assigns a premium to a risk  $X$  in terms of a functional of its distribution function.
- In practice, however, to determine an appropriate premium principle is only the first step in the process of risk pricing, because the involved distributions are not known so that  $H(X)$  can only be estimated by using the insured’s claim experience with certain effective statistical methods. These procedures form the so-called *experience rating*.
- We are concerned with solutions to this problem under the Bayesian setting: the distribution of a risk  $X$  is identified by an unknown and unobservable parameter (vector)  $\theta$ , via  $S(x, \theta) = \Pr(X > x|\theta)$ .

- The unobservable parameter  $\theta$  is modeled by a random vector. Its distribution  $\pi(\theta)$  is referred to as the *prior distribution* in statistics or *structural function* in actuarial science.
- The majority of the literature has mainly focussed on fully specified  $S(x, \theta)$  given  $\theta$  and known (or assumed)  $\pi(\theta)$ . Thus the problem can be solved with the standard parametric Bayesian methodology through posterior update.
- Although the applications of the parametric Bayesian methodology have been extensively investigated for general insurance, the reality is that  $\pi(\theta)$  is rarely known, making these applications impractical.
- To deal with such situations, one approach is to allow some unknown parameters in  $\pi(\theta)$ , but maintain some assumed mathematical forms of  $S(x, \theta)$  and  $\pi(\theta)$ . Then the risk pricing can be carried out by the empirical Bayes analysis.
- In many insurance practices, however, the mathematical forms of  $S(x, \theta)$  and  $\pi(\theta)$  are difficult to specify or unavailable. Under such circumstances, the analysis can only be done on a distribution-free or nonparametric basis.
- The best-known solution in the actuarial community is probably the Bühlmann's approach to net premium principle (known as the *credibility theory*).

- Under the net premium principle, the *credibility premium* of a future claim  $X_{n+1}$  can be expressed as a weighted average between the empirical individual mean (indicating the information delivered by historical data of the risk itself) and the collective mean (indicating the aggregated information obtained from all possible insureds):

$$\hat{H}_{n+1} = z\bar{X} + (1 - z)\mu_0, \quad (1.1)$$

where  $\bar{X} = n^{-1}\sum_{i=1}^n X_i$  is the empirical mean of the claims data  $\{X_1, X_2, \dots, X_n\}$ ,  $\mu_0 = E[X]$  is known as the *collective mean* of the risk, and  $z$  is the *credibility factor*.

- This (linear) weighted average has been so well received today that it is almost regarded as a synonym of the credibility theory.
- However, it is not easy to directly transplant Bühlmann’s method to other premium principles. Consequently most contributions in this area have been devoted to the net premium principle.
- Another important solution stream to this problem, originated from Ferguson (1973)<sup>1</sup>, is the *Bayesian nonparametrics*.
- Consider  $S(x, \theta)$  as a Dirichlet process with “time”  $x$ . Its distribution is determined by a finite Borel measure  $\alpha(\cdot)$  on the real line  $\mathbb{R}$ . After losses  $X_1, \dots, X_n$  are observed, the posterior distribution of  $\theta$  is also a Dirichlet process but with  $\alpha$  being updated to  $\alpha + \sum_{i=1}^n \delta_{X_i}$ , where  $\delta_{X_i}$  is a probability measure degenerated at  $X_i$ .

---

<sup>1</sup>Ferguson, T. A Bayesian analysis of some non-parametric problems. *Annals of Statistics*, **1**, 209-230.

- Consequently,  $S(x, \theta)$  can be estimated by the posterior mean

$$\widehat{S}(x, \theta) = \frac{\alpha(\mathbb{R})}{\alpha(\mathbb{R}) + n} \alpha([x, \infty)) + \frac{n}{\alpha(\mathbb{R}) + n} S_n(x), \quad (1.2)$$

where  $S_n(x) = n^{-1} \sum_{i=1}^n I(X_i > x)$ .

- Then the experience rating can be carried out by inserting the estimated distribution of the loss  $X$  into any premium principle.
- While the approach of Ferguson (1973) provides a unified solution for all premium principles, it requires the assumption of Dirichlet prior. This assumption is mainly for mathematical convenience, but hardly justifiable in practice.
- Without the Dirichlet assumption, the estimator in (1.2) would lose its justification, and the corresponding risk premium  $H(X_{n+1})$  would not be credible when the true prior deviates far from the Dirichlet.
- In this talk, we will introduce a new distribution-free approach for experience rating under arbitrary premium principles without any assumption on the prior distribution.
- It has two direct advantages:
  - (i) truly distribution-free settings as in the Bühlman's credibility theory, and
  - (ii) generating experience rating for arbitrary premium principles as achieved by Ferguson's Bayesian nonparametric method.

- The main idea is to first derive an estimate of  $S(x, \theta)$  by minimizing the  $L_2$ -distance, and then embed its estimate, say  $\widehat{S(x, \theta)}$ , into the premium principle functional to obtain an estimate of the corresponding premium, i.e.,  $\widehat{H(X, \theta)} = H(\widehat{S(x, \theta)})$ .

- Specifically, we consider the following two models:

**Model I:** The first two moments of  $S(x, \theta)$  with respect to the prior can be specified. In this case, the data are from  $X_1, X_2, \dots$ , which are (conditionally) i.i.d. copies of  $X$  given  $\theta$ . This model brings insight into and motivates the estimation methods of the more practically meaningful model II next.

**Model II:** This is a general model with totally unknown  $S(x, \theta)$ . Suppose that we have a portfolio of risks  $X_i, i = 1, 2, \dots, K$ , where each  $X_i$  is identified with a parameter (vector)  $\theta_i$  and contributes a series of data  $X_{i1}, X_{i2}, \dots, X_{in_i}, \dots$ , which are i.i.d. copies of  $X_i$ , given  $\theta_i$ . The purpose is to estimate the distributions and then assign proper premiums to the future claims  $X_{i, n_i+1}, i = 1, 2, \dots, K$  under a given premium principle.

- In order to overcome the restriction of Dirichlet priors, we develop a new approach in a quite different way from Ferguson (1973). It turns out, however, when the prior distribution is Dirichlet, our estimation happens to coincide with that of Ferguson (1973). This shows that our results actually generalize those of Ferguson (1973).

## 2 The Linear Bayes Estimation of Distributions

We consider Model I in this section. Write

$$S_0(x) = E_\theta [S(x, \theta)], \quad \tau_0^2(x) = \text{Var}_\theta [S(x, \theta)] \quad (2.1)$$

and

$$\sigma_0^2(x) = E_\theta [\text{Var}(I_{(X>x)}|\theta)], \quad (2.2)$$

where  $E_\theta$  and  $\text{Var}_\theta$  indicate that the expectations are computed with respect to the distribution of  $\theta$ . We also denote

$$\mu(\theta) = E[X|\theta], \quad \sigma^2 = E_\theta[\text{Var}(X|\theta)], \quad \tau^2 = \text{Var}_\theta(E(X|\theta)).$$

### The performance measure to be optimized

- As  $S(x, \theta) = E[I(X > x)|\theta]$ , a pointwise credibility estimator of  $S(x, \theta)$  for a fixed  $x$  following Bühlmann's classical credibility theory can be given by

$$S(x, \theta)^* = Z(x)S_n(x) + (1 - Z(x)) S_0(x), \quad (2.3)$$

where  $S_n(x) = n^{-1} \sum_{i=1}^n I(X_i > x)$  is the empirical estimate of  $S(x, \theta)$  based on the data  $X_1, X_2, \dots, X_n$ , and

$$Z(x) = \frac{n\tau_0^2(x)}{\sigma_0^2(x) + n\tau_0^2(x)}$$

is the credibility factor. By (2.1)–(2.2),

$$\sigma_0^2(x) = S_0(x) - E_\theta [S^2(x, \theta)], \quad \tau_0^2(x) = E_\theta [S^2(x, \theta)] - S_0^2(x). \quad (2.4)$$

- The naive estimate in (2.3) can be more precisely expressed as

$$\begin{aligned}
S(x, \theta)^* &= \frac{n(\mathbb{E}_\theta[S^2(x, \theta)] - S_0^2(x))S_n(x) + (S_0(x) - \mathbb{E}_\theta[S^2(x, \theta)])S_0(x)}{n(\mathbb{E}_\theta[S^2(x, \theta)] - S_0^2(x)) + (S_0(x) - \mathbb{E}_\theta[S^2(x, \theta)])} \\
&= \sum_{i=1}^{n+1} \frac{(n-i+1)\mathbb{E}_\theta[S^2(x, \theta)] - (n-i)S_0^2(x) - S_0(x)\mathbb{E}_\theta[S^2(x, \theta)]}{(n-1)\mathbb{E}_\theta[S^2(x, \theta)] - nS_0^2(x) + S_0(x)} I_{[X_{(i-1)}, X_{(i)}]}(x),
\end{aligned} \tag{2.5}$$

where  $-\infty = X_{(0)} < X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)} < X_{(n+1)} = \infty$  are the order statistics of the sample.

- This naive credibility estimate (2.3) (or (2.5)) would require:
  - checking monotonicity of  $S(x, \theta)^*$  in  $x$ ,
  - making an intuitive interpretation for  $S(x, \theta)^*$ , and
  - computing a premium under  $S(x, \theta)^*$ .

These tasks are however not feasible due to the dependence of the credibility factor  $Z(x)$  on  $x$ .

- A natural remedy is to consider  $Z(x) = Z$  to be independent of  $x$ . More specifically, we seek a function of the form

$$\beta_0(x) + \sum_{i=1}^n \lambda_i I(X_i > x), \tag{2.6}$$

which minimizes the square of the  $L_2$ -distance:

$$\min_{\beta_0(x), \lambda_1, \dots, \lambda_n} \mathbb{E} \left[ \int_{-\infty}^{+\infty} \left( S(x, \theta) - \beta_0(x) - \sum_{i=1}^n \lambda_i I(X_i > x) \right)^2 dx \right], \tag{2.7}$$

where  $\beta_0(x)$  is a nonincreasing function independent of the claims history and  $\lambda_1, \dots, \lambda_n$  are real-valued decision variables.



## Estimation of distribution functions

The following theorem gives an estimator of  $S(x, \theta)$  by solving the optimization problem (2.7).

**Theorem 2.1** *The credibility estimator of  $S(x, \theta)$ , which minimizes (2.7), is given by*

$$\widehat{S(x, \theta)} = Z S_n(x) + (1 - Z) S_0(x), \quad (2.8)$$

where

$$Z = \frac{n\tau_0^2}{n\tau_0^2 + \sigma_0^2} \quad (2.9)$$

with

$$\tau_0^2 = \int_{-\infty}^{\infty} \tau_0^2(x) dx \quad \text{and} \quad \sigma_0^2 = \int_{-\infty}^{\infty} \sigma_0^2(x) dx.$$

Moreover, the integrated mean-squared loss for the optimal estimator given by (2.8) is

$$\int_{-\infty}^{+\infty} \text{E} \left( \widehat{S(x, \theta)} - S(x, \theta) \right)^2 dx = \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2}. \quad (2.10)$$

Note that Theorem 2.1 does not require to specify the prior distribution. If the prior is not Dirichlet, then Ferguson's method cannot be applied, but the estimator defined in (2.8) and (2.9) remains valid. In this sense, our new estimation provides a more general solution for the estimation of the distributions than that of Ferguson (1973).

## Properties of estimator

- If  $S(x, \theta)$  is a Dirichlet process with a finite Borel measure  $\alpha(\cdot)$  on  $\mathbb{R}$ , then

$$S(x, \theta) \sim \text{Beta}(\alpha(x), \alpha(\mathbb{R}) - \alpha(x))$$

with density

$$\frac{\Gamma(\alpha(\mathbb{R}))}{\Gamma(\alpha(x))\Gamma(\alpha(\mathbb{R}) - \alpha(x))} t^{\alpha(x)-1} (1-t)^{\alpha(\mathbb{R})-\alpha(x)-1},$$

where  $\alpha(x) = \alpha((x, \infty))$ .

- It is easy to show that in such a case, the credibility factor  $Z$  in (2.9) becomes

$$Z = \frac{n\tau_0^2}{n\tau_0^2 + \sigma_0^2} = \frac{\alpha(\mathbb{R})}{\alpha(\mathbb{R}) + n}.$$

Thus the estimate in (2.8) reduces to Ferguson's estimate (1.2).

- The estimator in (2.8) is obviously a survival distribution function.
- By (2.9),  $0 \leq Z \leq 1$ ,

$$Z \rightarrow 1 \text{ as } n \rightarrow \infty \quad \text{and} \quad Z \rightarrow 0 \text{ as } n \rightarrow 0, \quad (2.11)$$

which allows the traditional credibility interpretation: the more data, the more credible the empirical distribution  $S_n(x)$ .

- The estimator (2.8) is uniformly strongly consistent in the sense that

$$\lim_{n \rightarrow \infty} \sup_x |S(x, \theta) - \widehat{S}(x, \theta)| = 0 \quad \text{almost surely (a.s.).}$$

### 3 The Empirical Bayes Estimation

We now consider Model II, in which both the conditional distribution  $S(x, \theta)$  and the structure function  $\pi(\theta)$  are completely unknown, including the structure parameters  $S_0(x)$ ,  $\tau_0^2$  and  $\sigma_0^2$ .

- We propose an empirical Bayes method to estimate  $S_0(x)$ ,  $\tau_0^2$  and  $\sigma_0^2$  based on the claim experiences over a number of risks in the same portfolio.
- Let  $X_1, \dots, X_K$  denote  $K$  risks under observation. The distribution of each  $X_i$  is characterized by a risk parameter  $\theta_i$  and contributes a sequence of claim experiences  $X_i = (X_{i1}, \dots, X_{in_i})$  over  $n_i$  time periods, subject to the following usual assumptions.

- **Assumption 3.1**

(i) *Conditional on  $\theta_i$ , the random variables  $X_{ij}$  ( $j = 1, 2, \dots, n_i$ ) are i.i.d. with common unknown distribution and moments:*

$$S(x, \theta_i) = \Pr(X_{ij} > x | \theta_i), \quad \sigma^2(x, \theta_i) = \text{Var}\left(I_{(X_{ij} > x)} | \theta_i\right),$$

$$i = 1, 2, \dots, K, \quad j = 1, 2, \dots, n_i.$$

(ii) *The random variables  $\theta_1, \dots, \theta_K$  are i.i.d. with a common prior distribution  $\pi(\theta)$ .*

- Write  $S_0(x) = E_\theta [S(x, \theta_i)]$ ,  $\tau_0^2(x) = \text{Var}_\theta [S(x, \theta_i)]$  and  $\sigma_0^2(x) = E_\theta [\sigma^2(x, \theta_i)]$ .

## Homogeneous estimation of the distribution

- Write

$$S_i(x) = \frac{1}{n_i} \sum_{j=1}^{n_i} I(X_{ij} > x), \quad i = 1, 2, \dots, K. \quad (3.1)$$

- To obtain the credibility estimator of  $S(x, \theta_i)$ , consider the class of distribution functions

$$L = \left\{ \sum_{s=1}^K \sum_{t=1}^{n_s} a_{st} I(X_{st} > x), a_{st} \in \mathbb{R}, \sum_{s=1}^K \sum_{t=1}^{n_s} a_{st} = 1 \right\}, \quad (3.2)$$

- Solve the minimization problem

$$\min_{g \in L} \mathbb{E} \left\{ \int_{-\infty}^{+\infty} [g - S(x, \theta_i)]^2 dx \right\}. \quad (3.3)$$

The solution is stated in the next theorem.

**Theorem 3.1** *The homogeneous credibility estimator of  $S(x, \theta_i)$  as the solution to (3.3) is*

$$\widehat{S(x, \theta_i)}^* = Z_i S_i(x) + (1 - Z_i) \widetilde{S}(x), \quad , i = 1, \dots, K, \quad (3.4)$$

where

$$Z_i = \frac{n_i \tau_0^2}{\sigma_0^2 + n_i \tau_0^2} \quad \text{and} \quad \widetilde{S}(x) = \frac{1}{\sum_{i=1}^K Z_i} \sum_{i=1}^K Z_i S_i(x). \quad (3.5)$$

We can also derive an *inhomogeneous* estimator of the form

$$\beta_0(x) + \sum_{s=1}^K \sum_{t=1}^{n_s} a_{st} I(X_{st} > x).$$

The result, however, is the same as that presented by (2.8)–(2.9). This is a typical phenomenon in credibility theory.

## Estimation of structure parameters

- We now proceed to the estimation of the structure parameters  $\tau_0^2$  and  $\sigma_0^2$ , which needs the estimation of  $S_0(x)$ .
- At this moment, however,  $\widetilde{S}(x)$  is not a suitable estimate of  $S_0(x)$  because  $Z_i$ 's involve  $\tau_0^2$  and  $\sigma_0^2$  (cf. (3.5)).
- Moreover, if we use an estimator of the form  $\overline{S}(x) = \sum_{i=1}^K w_i S_i(x)$  by minimizing  $\int (\overline{S}(x) - S_0(x))^2 dx$ , it leads to  $\overline{S}(x) = \widetilde{S}(x)$ .
- Therefore, we simply select

$$\overline{S}(x) = \frac{1}{N} \sum_{i=1}^K n_i S_i(x), \quad \text{where} \quad N = \sum_{i=1}^K n_i, \quad (3.6)$$

which is unbiased and strongly consistent in the sense that  $\overline{S}(x)$  converges to  $S_0(x)$  almost surely as  $K \rightarrow \infty$ , under the condition

$$\sum_{K=1}^{\infty} \frac{n_K^2}{(N - K)^2} < \infty. \quad (3.7)$$

- In view of the definitions of  $\sigma_0^2(x)$  and  $\tau_0^2(x)$ , the parameters  $\sigma_0^2$  and  $\tau_0^2$  can be estimated using

$$SSE(x) = \sum_{i=1}^K \sum_{j=1}^{n_i} [I(X_{ij} > x) - S_i(x)]^2$$

and

$$SSA(x) = \sum_{i=1}^K n_i \left[ S_i(x) - \overline{S}(x) \right]^2.$$

- The estimators can be formally defined as

$$\widehat{\sigma}_0^2 = \frac{1}{N - K} \int_{-\infty}^{\infty} SSE(x) dx \quad (3.8)$$

and

$$\widehat{\tau}_0^2 = \frac{N}{N^2 - \sum_{i=1}^K n_i^2} \left[ \int_{-\infty}^{\infty} SSA(x) dx - \frac{K-1}{N-K} \int_{-\infty}^{\infty} SSE(x) dx \right]. \quad (3.9)$$

- By inserting the estimates of  $\tau_0^2$  and  $\sigma_0^2$  into (3.4) and (3.5), we can get *empirical Bayes estimators* for the distribution  $S(x, \theta_i)$  as

$$\widetilde{S}(x, \theta_i) = \widehat{Z}_i S_i(x) + (1 - \widehat{Z}_i) \widehat{S}_0(x), \quad i = 1, \dots, K, \quad (3.10)$$

where

$$\widehat{Z}_i = \frac{n_i \widehat{\tau}_0^2}{\widehat{\sigma}_0^2 + n_i \widehat{\tau}_0^2}, \quad i = 1, \dots, K, \quad (3.11)$$

and

$$\widehat{S}_0(x) = \frac{1}{\sum_{i=1}^K \widehat{Z}_i} \sum_{i=1}^K \widehat{Z}_i S_i(x). \quad (3.12)$$

- The properties of the estimators in (3.8)–(3.9) are provided in the following theorem.

**Theorem 3.2**  $\widehat{\sigma}_0^2$  and  $\widehat{\tau}_0^2$  have the following properties.

- (1)  $\widehat{\sigma}_0^2$  and  $\widehat{\tau}_0^2$  are unbiased estimators of  $\sigma_0^2$  and  $\tau_0^2$ , respectively.
- (2) Under condition (3.7),  $\widehat{\sigma}_0^2 \rightarrow \sigma_0^2$  and  $\widehat{\tau}_0^2 \rightarrow \tau_0^2$  almost surely as  $K \rightarrow \infty$ .

## Asymptotic optimality

- For simplicity, we discuss the balanced case with  $n_i \equiv n$ , thus  $Z_i$  and  $\widehat{Z}_i$  are all the same. Consequently,

$$\widehat{S}_0(x) = \frac{1}{K} \sum_{i=1}^K S_i(x).$$

- An additional condition is needed for the asymptotic optimality needs the following extra: there is a constant  $\delta > 0$  such that

$$\int_{-\infty}^{+\infty} \left( \frac{\sigma_0^2(x)}{n} + \tau_0^2(x) \right)^{\delta/(2+\delta)} dx < \infty. \quad (3.13)$$

- Write

$$\widetilde{Q}(x, \theta_i) = \int_{-\infty}^{+\infty} \left( \widetilde{S}(x, \theta_i) - S(x, \theta_i) \right)^2 dx$$

and

$$\widehat{Q}(x, \theta_i) = \int_{-\infty}^{+\infty} \left( \widehat{S}(x, \theta_i) - S(x, \theta_i) \right)^2 dx.$$

- The asymptotic optimality is stated in the theorem below.

**Theorem 3.3** *Under conditions (3.7) and (3.13), the  $\widetilde{S}(x, \theta_i)$  in (3.10)–(3.12) are asymptotically optimal in the sense that*

$$\lim_{K \rightarrow \infty} \max_{1 \leq i \leq K} \left| \mathbb{E} \left[ \widetilde{Q}(x, \theta_i) - \widehat{Q}(x, \theta_i) \right] \right| = 0, \quad (3.14)$$

where  $\widehat{S}(x, \theta_i)$  is the linear estimator defined by (2.8) and (2.9).

## 4 Applications to Experience Ratemaking

- Having estimated the distributions, the application to the experience ratemaking is apparent.
- For a given premium principle  $H$ , write  $H(X|\theta)$  for the individual premium of a risk  $X$  with distribution  $S(x, \theta)$ , referred also to as the *risk premium* of  $X$ .
- By replacing  $S(x, \theta)$  with  $\widehat{S}(x, \theta)$  or  $\widetilde{S}(x, \theta_i)$ , an estimator  $\widehat{H}(X|\theta)$  of the individual premium  $H(X|\theta)$  is obtained.
- We consider the experience premium by inserting  $\widehat{S}(x, \theta)$  for  $S(x, \theta)$  and compare it with existing methods. The version using  $\widetilde{S}(x, \theta_i)$  can be discussed in a similar way.
- The consistency of  $\widehat{H}(X|\theta)$  is presented in the theorem below, which follows directly from Theorems 2.1 and 3.2.

**Theorem 4.1** *If the premium principle  $H$  is continuous with respect to the  $L^\infty$ -norm of distribution functions, then  $\widehat{H}(X|\theta)$  is strongly consistent for  $H(X|\theta)$ .*

- According to Theorem 4.1, the strong consistency of  $\widehat{H}(X|\theta)$  for  $H(X|\theta)$  is guaranteed as long as the premium principle  $H$  satisfies the condition of the theorem.



- This is in contrast to existing results in the literature, where either the credibility estimator is not generally (strongly) consistent, or the consistency needs to be proved separately in each case.
- There are a number of well known premium principles for which strong consistency of the experience ratemaking can be obtained directly or by checking the condition of Theorem 4.1, including

- the net premium  $H(X|\theta) = E[X|\theta]$ ,

- variance premium  $H(X|\theta) = E[X|\theta] + \alpha \text{Var}H(X|\theta)$ ,

- modified variance premium

$$H(X|\theta) = E[X|\theta] + \alpha \frac{\text{Var}(X|\theta)}{E[X|\theta]},$$

- standard deviation premium

$$H(X|\theta) = E[X|\theta] + \alpha \sqrt{\text{Var}(X|\theta)},$$

- Esscher premiums

$$H(X|\theta) = \frac{E[Xe^{hX}|\theta]}{E[e^{hX}|\theta]},$$

- conditional tail expectation premium  $H(X|\theta) = E[X|X > \alpha, \theta]$ ,

- exponential premium  $H(X|\theta) = \alpha^{-1} \log [E(e^{\alpha X}|\theta)]$ , and

- Kamp's premium

$$H(X|\theta) = \frac{E[X(1 - e^{\alpha X})|\theta]}{E[(1 - e^{\alpha X})|\theta]}.$$

- The following are two exceptions where the proof of the consistency require additional work and conditions.

1. *Dutch's premium principle:*

$$H(X|\theta) = E[X|\theta] + \eta E[(X - \alpha E[X|\theta])_+ | \theta],$$

where  $\alpha \geq 1$  and  $0 < \eta \leq 1$ . Observe that

$$E_{S_n}[(X - \alpha E[X|\theta])_+] = \frac{1}{n} \sum_{i=1}^n (X_i - \alpha \widehat{\mu}(\theta))_+.$$

The estimate of  $H(X|\theta)$  is then given by

$$\widehat{H}(X|\theta) = \widehat{\mu}(\theta) + \eta \left\{ \frac{Z}{n} \sum_{i=1}^n (X_i - \alpha \widehat{\mu}(\theta))_+ + (1 - Z) \int_{-\infty}^{\infty} (x - \alpha \widehat{\mu}(\theta))_+ dF_0(x) \right\},$$

where  $\widehat{\mu}(\theta) = Z \overline{X} + (1 - Z)\mu_0$ .

2. *Distortion premium principle:*

$$H(X|\theta) = \int_0^{\infty} g(S(x, \theta)) dx - \int_{-\infty}^0 g(1 - S(x, \theta)) dx.$$

The premium estimate is

$$\widehat{H}(X|\theta) = \int_0^{\infty} g(Z S_n(x) + (1 - Z) S_0(x)) dx - \int_{-\infty}^0 g(Z F_n(x) + (1 - Z) F_0(x)) dx,$$

where the credibility factor is given by (2.9).

The consistency in this case can be proved under certain additional regularity condition on  $g$ , such as the Lipschitz condition of  $g$  on  $[0, 1]$  if  $X$  has bounded support, or of  $g'$  on  $[0, 1]$  if the support is unbounded.

## 5 Performance Evaluation

We evaluate the performance of the proposed approach by comparing it with some well-known representative methods in the literature.

### Evaluation under the net premium principle

- Under the net premium principle, the optimal premium is given by the Bühlmann's credibility premium

$$\widehat{\mu(\theta)}^c = Z^c \bar{X}_n + (1 - Z^c) \mu_0,$$

where the superscript “c” stands for “classical”.

- The premium using our method is

$$\widehat{\mu(\theta)} = \widehat{H(X|\theta)} = Z \bar{X} + (1 - Z) \mu_0. \quad (5.1)$$

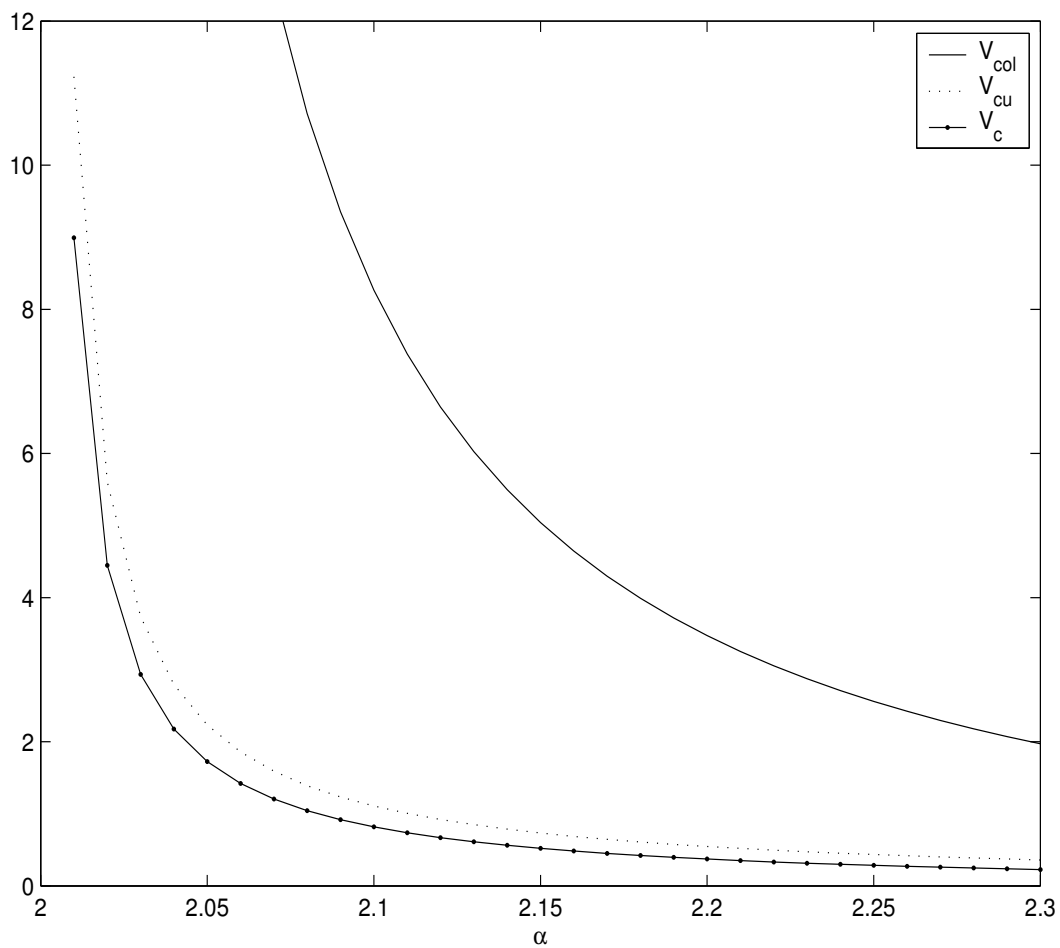
- The expected quadratic losses of  $\widehat{\mu(\theta)}$  approaches that of  $\widehat{\mu(\theta)}^c$ :

$$\lim_{n \rightarrow \infty} \frac{\text{E}[\widehat{\mu(\theta)} - \mu(\theta)]^2}{\text{E}[\widehat{\mu(\theta)}^c - \mu(\theta)]^2} = 1. \quad (5.2)$$

This shows that  $\widehat{\mu(\theta)}$  is asymptotically equivalent to the optimum.

- Assume  $S(x, \theta) = e^{-\theta x}$  and  $\theta \sim \text{Gamma}(\alpha, \beta)$ . Then the expected squared losses of  $\widehat{\mu(\theta)}$ ,  $\widehat{\mu(\theta)}^c$  and  $\mu_0$  (the collective premium) can be expressed by  $\beta^2 V_{\text{cu}}(\alpha)$ ,  $\beta^2 V_{\text{c}}(\alpha)$  and  $\beta^2 V_{\text{col}}$  respectively.
- The curves of  $V_{\text{cu}}(\alpha)$ ,  $V_{\text{c}}(\alpha)$  and  $V_{\text{col}}(\alpha)$  as functions of  $\alpha$  for  $n = 10$  are plotted in Figure 1 below, which shows that  $\widehat{\mu(\theta)}$  is close to the optimal  $\widehat{\mu(\theta)}^c$ .

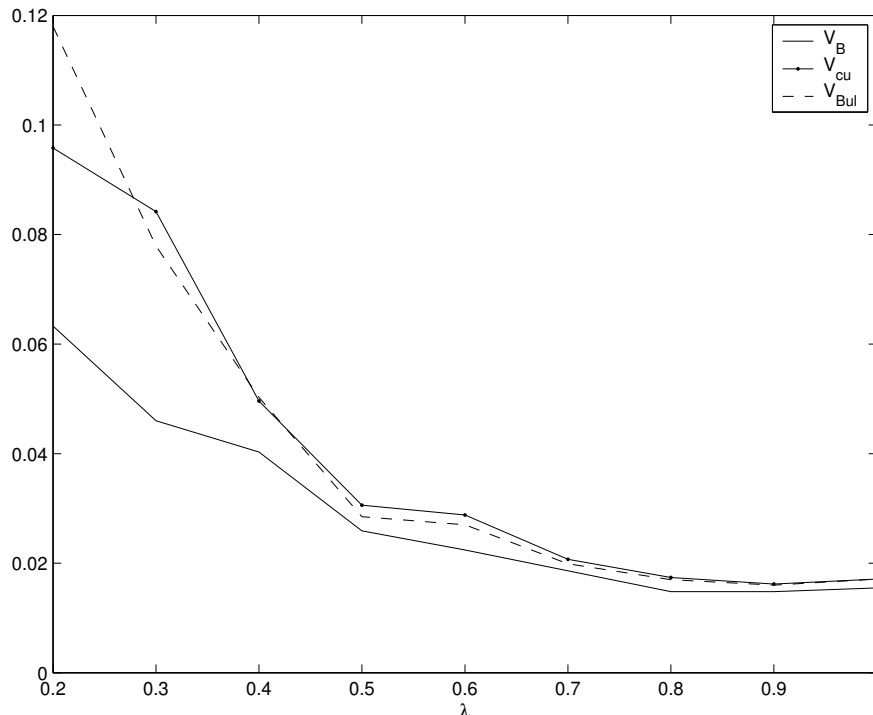
Figure 1. The curves of  $V_{cu}(\alpha)$ ,  $V_c(\alpha)$  and  $V_{col}(\alpha)$  as functions of  $\alpha$ .



## Evaluation under the variance premium principle

- Consider the Poisson-exponential model, where  $S(x, \theta)$  is Poisson with mean  $\theta$  and  $\pi(\theta)$  is exponential with rate  $\lambda$ .
- Bühlmann suggested a credibility premium formula that involves the fourth moments of  $X_i$  given  $\theta$ . In comparison, we only need the first two moments of the risk distribution.
- Let  $V_{\&}$  denote the expected squared loss with  $\& = B, cu, Bul, col$  corresponding to Bayes, proposed, Bühlmann, collective premiums respectively under the variance premium principle.
- Figure 2 below shows the numerical results of  $V_{\&}$  for  $\lambda = 0.2$  to  $1.0$  with sample size  $n = 100$ .

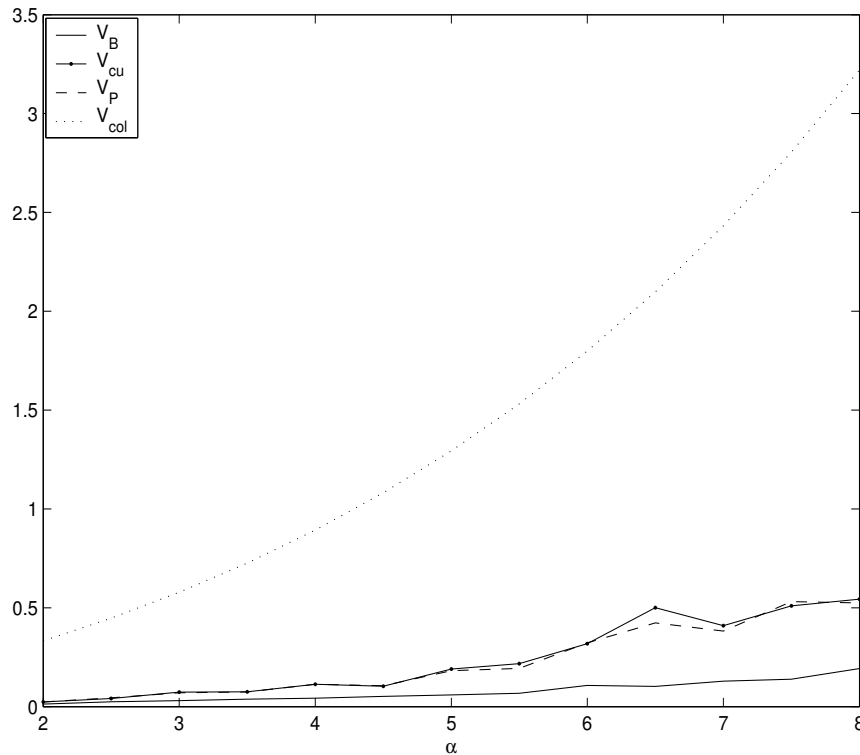
Figure 2. The curves of  $V_B, V_{cu}$  and  $V_{Bul}$  as functions of  $\lambda$  with  $n = 100$



## Evaluation under the Esscher premium principle

- Consider the Poisson-Gamma model with  $S(x, \theta) \sim \text{Poisson}(\theta)$  and  $\theta \sim \text{Gamma}(\alpha, \beta)$  with  $\alpha$  ranging from 2 to 8 and  $\beta = 6$ .
- Let  $V_{\&}$  denote the expected squared loss with  $\& = \text{B, cu, col, P}$  under the Esscher premium principle, where “P” denotes the method introduced in Pan *et.al* (2008)<sup>2</sup>.
- Figure 3 below shows the numerical results of  $V_{\&}$  for different values of  $\alpha$  and sample size 100.

Figure 3. Curves of  $V_{\&}$  for different values of  $\alpha$  and sample size 100



---

<sup>2</sup>Pan, M., Wang, R. and Wu, X. On the consistency of credibility premiums regarding Esscher principle. *Insurance: Mathematics and Economics*, **42**, 119-126, 2008.