应用概率统计　第 38 卷
第 2 期　2022 年 4 月

Chinese Journal of Applied Probability and Statistics
Apr., 2022, Vol. 38, No. 2, pp. 237-252
doi: 10.3969/j.issn.1001-4268.2022.02.005

二元贝叶斯聚合风险模型中保费的后验厘定 *

章　溢

(江西师范大学财政金融学院, 南昌, 330022;　江西财经大学金融学院, 南昌, 330013)

摘　要: 将索赔额区分为大额索赔和小额索赔, 在方差相关保费原理下研究了二元贝叶斯聚合风险模型中风险保费的贝叶斯估计. 结论显示, 风险的条件期望和条件方差都能表达为样本函数和聚合保费的加权形式, 其中权重满足 "信度因子" 的性质. 进而, 证明了贝叶斯估计的强相合性和渐近正态性. 最后, 利用数值模拟的方法验证了贝叶斯估计的大样本性质.

关键词: 聚合风险模型; 风险保费; 方差相关保费原理; 信度估计; 渐近正态性

中图分类号: O211.9

英文引用格式: ZHANG Y. The posterior ratemaking of premium in binary Bayesian collective risk model [J]. Chinese J Appl Probab Statist, 2022, 38(2): 237–252. (in Chinese)

§1. 引　　言

在汽车保险中, 常用的奖惩系统都只是基于索赔次数的一种奖惩机制. 然而, 同一次索赔中, 对索赔额小的事故与索赔额大的事故进行相同的惩罚显然是不公平的. 因此, 更好的奖惩系统不仅需要考虑索赔次数, 也应该考虑索赔额的大小. Lemaire [1] 指出, 只有当索赔额与索赔次数相互独立的时候, 仅考虑索赔次数的奖惩系统才是合理的. Frangos 和 Vrontos [2] 提出汽车保险的频率和额度的二元奖惩系统设计方法, Gómez-Déniz 等 [3] 利用信度理论方法对总索赔分布的奖惩保费设计进行了分析, Gómez-Déniz [4] 对区分两种类型索赔的聚合风险模型给出了基于信度理论的奖惩系统设计. 研究结果和索赔数据的经验显示, 汽车保险中索赔额和索赔次数常常呈现正相依性. Liu 和 Wang [5] 研究了具有不确定性相依的聚合风险模型, Cossette 等 [6] 研究了聚合风险模型的一般相依结构. 近年来, 关于聚合风险模型中索赔次数和索赔额相依的研究是非寿险精算的热点问题, 相关研究可参考文献 [7–9] 等.

在对聚合风险模型的总风险制定保费时, 不仅需要考虑总风险的数学期望, 而且需要考虑总风险的方差. 由风险的数学期望或方差组合构成的保费原理包括期望值保费原理、方差保费原理、标准差保费原理和修正方差保费原理等 [10]. Bühlmann [11] 指出, 标准差原理是财产保险与意外事故保险中使用最多的保费原理, 而方差保费原理是理论研究广泛的

*国家自然科学基金项目 (批准号: 71761019) 与江西省教育厅科学技术研究项目 (批准号: GJJ200304) 资助.
E-mail: yizi85820@163.com.
本文 2020 年 9 月 7 日收到, 2020 年 10 月 30 日收到修改稿.

一种保费原理. 本文将在更一般的保费原理——方差相关保费原理中研究聚合风险的后验厘定问题. 方差相关保费原理是 Guerra 和 Centeno[12] 在最优再保险的研究中首次提出的, 相关的研究可参考文献 [13,14] 等. 本文后面的内容安排如下. 第二节建立区分额度索赔的贝叶斯聚合风险模型, 第三节研究方差相关保费原理中风险保费的贝叶斯估计, 并给出估计的统计性质. 在第四节中, 利用数值模拟的方法验证估计的统计性质, 进一步说明本文估计的优势.

§2. 区分额度索赔的聚合风险模型

设 N 表示该保单期内的索赔次数, 为取非负整数值的离散型随机变量, 而 $\{Y_i, i \geqslant 1\}$ 为每次的索赔额, 则总风险为 $Y_1 + Y_2 + \cdots + Y_N$. 称该模型为聚合风险模型[14]. 在传统的聚合风险模型中, 常常假设索赔次数和索赔额相互独立, 且每次的索赔额 $\{Y_i, i \geqslant 1\}$ 相互独立并具有共同的分布. 在实际的运用中, 索赔次数和索赔额常常呈现一定的相依性, 例如一次大的交通事故常常导致索赔次数和索赔额有正的相依性. 为了刻画这种相依性, 本文将建立区分大额索赔和小额索赔的贝叶斯二元聚合风险模型.

在二元聚合风险模型中, 根据某个临界值 d 将索赔额 $\{Y_i, i \geqslant 1\}$ 分成两类索赔: 若 $Y_i < d$, 则称该索赔为小额索赔, 若 $Y_i \geqslant d$, 则称该索赔为大额索赔, 这里 d 为已知的确定的实数[4]. 令 $X_i = I(Y_i \geqslant d)$ 与 $Z_i = I(Y_i < d)$ 表示额度索赔的示性变量, 则 $X_i + Z_i = 1$ 且 X_i 和 Z_i 服从两点分布随机变量. 记 $Z = \sum\limits_{i=1}^{N} Z_i$ 和 $X = \sum\limits_{i=1}^{N} X_i$ 分别表示所有索赔中小额索赔和大额索赔的索赔次数. 本文将在二元贝叶斯聚合风险模型中探讨 (X, Z) 的联合分布和函数 $h(X, Z)$ 的预测问题. 因此, 给定下面的假设.

假设 1 假定在 θ_1 给定条件下, 索赔次数 $N_1, N_2, \cdots, N_m, N_{m+1}$ 相互独立且服从泊松分布, 其条件概率函数为

$$\mathsf{P}(N_i = k \,|\, \theta_1) = \frac{\theta_1^k}{k!}\mathrm{e}^{-\theta_1}, \qquad k = 0, 1, 2, \cdots. \tag{1}$$

假设 2 假定在 θ_2 给定条件下, 二元随机向量 $\{(X_{ij}, Z_{ij}), i \geqslant 1, j \geqslant 1\}$ 相互独立且服从相同的分布, 其边际分布为两点分布:

$$\mathsf{P}(Z_{ij} = 1 \,|\, \theta_2) = \theta_2 = 1 - \mathsf{P}(Z_{ij} = 0 \,|\, \theta_2) \ \text{且}\ X_{ij} + Z_{ij} = 1. \tag{2}$$

记

$$X_i = \sum_{j=1}^{N_i} X_{ij} \quad \text{以及} \quad Z_i = \sum_{j=1}^{N_i} Z_{ij} \tag{3}$$

分别表示第 i 个保单的大额索赔和小额索赔次数总数, 其中 $i = 1, 2, \cdots, m, m+1$. 这里 m 表示可观测到的样本容量.

假设 3 在风险参数 $\theta=(\theta_1,\theta_2)$ 给定条件下, 索赔次数 $\{N_1,N_2,\cdots,N_m,N_{m+1}\}$ 与索赔示性变量 $\{X_{ij},Z_{ij},i\geqslant 1,j\geqslant 1\}$ 相互独立, 且风险参数 θ_1 和 θ_2 为相互独立的随机变量, 边际先验分布密度分别为

$$\pi_1(\theta_1)=\frac{\beta^\alpha}{\Gamma(\alpha)}\theta_1^{\alpha-1}\mathrm{e}^{-\beta\theta_1}I(\theta_1>0) \tag{4}$$

以及

$$\pi_2(\theta_2)=\frac{1}{\mathrm{Be}(a,b)}\theta_2^{a-1}(1-\theta_2)^{b-1}I(0<\theta_2<1). \tag{5}$$

在假设 1–3 中, 记 $\widetilde{S}=\{N_1,\cdots,N_m,X_{11},\cdots,X_{m,N_m},Z_{11},\cdots,Z_{m,N_m}\}$ 表示所有样本, 其观测值记为 $\widetilde{s}=\{n_1,\cdots,n_m,x_{11},\cdots,x_{m,n_m},z_{11},\cdots,z_{m,n_m}\}$. 而

$$X_{m+1}=\sum_{j=1}^{N_{m+1}}X_{m+1,j}\quad 和\quad Z_{m+1}=\sum_{j=1}^{N_{m+1}}Z_{m+1,j}$$

为待预测的量. 更一般地, 需要预测 X_{m+1} 与 Z_{m+1} 的线性组合, 定义待预测的风险为

$$V=w_lX_{m+1}+w_sZ_{m+1},$$

其中 w_l 和 w_s 分别表示大额索赔和小额索赔的权重系数.

§3. 风险保费的贝叶斯估计

设 R 为风险随机变量, 定义

$$H(R)=\mathsf{E}R+g(\mathsf{E}R,\mathsf{Var}\,(R)), \tag{6}$$

称为风险 R 的方差相关保费原理, 其中 $g(x,y)$ 为连续的二元函数. 当 $g(x,y)$ 取为不同的函数时, 该保费原理退化为一些常用的保费原理. 例如, 当 $g(x,y)=\xi x$ 时为期望值保费原理; 当 $g(x,y)=\xi y$ 时为方差保费原理; 当 $g(x,y)=\xi\sqrt{y}$ 时为标准差保费原理, 等等. 关于保费原理的总结, 可参考文献 [10].

在寿险精算中, 精算师可以根据生命表较为精确地估计不同年龄的被保险人的生存概率和死亡概率. 因此, 净保费原理或期望值保费原理在寿险精算中使用较为广泛. 例如, 在年金的计算或寿险责任准备金的计算中, 常常根据净保费原理下的平衡原则计算年金的净值或责任准备金总额[15]. 而对于非寿险精算, 由于风险因素较多, 影响索赔额和索赔次数的变量较多, 需要更多地关注风险的二阶甚至更高阶矩. 因此, 同时基于风险的数学期望和方差的保费原理得到了更广泛的应用. 例如, 在财产保险和责任保险中, 精算师 Bühlmann 更推荐标准差保费原理定价保费[11]. 关于标准差保费原理、方差保费原理、修正方差保费原理的应用, 更多的可参考文献 [16–18]. 本文提出的方差相关保费原理是这几种保费原理的综合, 通过选取二元函数 $g(x,y)$, 得到不同的保费计算原理. 因此在实际中有更广泛的应用.

注记 4 文献 [12] 在最优再保险的研究中首次提出方差相关保费原理, 定义为

$$H(R) = \mathsf{E}R + h(\mathsf{Var}\,(R)), \tag{7}$$

其中 $h(x)$ 为一元连续函数. 然而, 该保费原理不能包含修正方差保费原理, 而 (6) 不仅可以包含修正方差保费原理, 而且当 $g(x,y) = h(x)$ 时退化为 (7) 的保费原理. 因此, 本文提出的相关保费原理是文献 [12] 的简单推广.

在二元贝叶斯聚合风险模型中, 风险 V 的条件期望和条件方差分别记为

$$\mu_v(\theta) = \mathsf{E}(V\,|\,\theta) \quad 和 \quad \sigma_v^2(\theta) = \mathsf{Var}\,(V\,|\,\theta).$$

在方差相关保费原理 (6) 中, 风险 V 的风险保费为

$$R_v(\theta) = \mu_v(\theta) + g(\mu_v(\theta), \sigma_v^2(\theta)).$$

引理 5 在假设 1–3 中, 若 $\theta = (\theta_1, \theta_2)$ 给定, 则风险 V 的条件期望和条件方差分别为

$$\mu_v(\theta) = (w_l - w_s)\theta_1\theta_2 + w_s\theta_1 \quad 以及 \quad \sigma_v^2(\theta) = (w_l^2 - w_s^2)\theta_1\theta_2 + w_s^2\theta_1. \tag{8}$$

因而得到方差相关保费原理中的风险保费为

$$R_v(\theta) = (w_l - w_s)\theta_1\theta_2 + w_s\theta_1 + g((w_l - w_s)\theta_1\theta_2 + w_s\theta_1, (w_l^2 - w_s^2)\theta_1\theta_2 + w_s^2\theta_1).$$

证明: 注意到 $X_{m+1} + Z_{m+1} = N_{m+1}$, 且在 N_{m+1} 和 θ 给定条件下 $(X_{m+1}\,|\,N_{m+1}, \theta)$ 与 $(Z_{m+1}\,|\,N_{m+1}, \theta)$ 分别服从二项分布 $\mathrm{B}(N_{m+1}, \theta_2)$ 和 $\mathrm{B}(N_{m+1}, 1-\theta_2)$. 因此有

$$\mathsf{E}(X_{m+1}\,|\,N_{m+1}, \theta) = N_{m+1}\theta_2, \qquad \mathsf{E}(Z_{m+1}\,|\,N_{m+1}, \theta) = N_{m+1}(1-\theta_2)$$

以及

$$\mathsf{Var}\,(X_{m+1}\,|\,N_{m+1}, \theta) = N_{m+1}\theta_2(1-\theta_2), \qquad \mathsf{Var}\,(Z_{m+1}\,|\,N_{m+1}, \theta) = N_{m+1}\theta_2(1-\theta_2).$$

根据双重期望公式有

$$\begin{aligned}
\mu_v(\theta) &= \mathsf{E}(w_l X_{m+1} + w_s Z_{m+1}\,|\,\theta) \\
&= w_l\mathsf{E}(X_{m+1}\,|\,\theta) + w_s\mathsf{E}(Z_{m+1}\,|\,\theta) \\
&= w_l\mathsf{E}[\mathsf{E}(X_{m+1}\,|\,\theta, N_{m+1})\,|\,\theta] + w_s\mathsf{E}[\mathsf{E}(Z_{m+1}\,|\,\theta, N_{m+1})\,|\,\theta] \\
&= w_l\mathsf{E}(N_{m+1}\theta_2\,|\,\theta) + w_s\mathsf{E}[N_{m+1}(1-\theta_2)\,|\,\theta] \\
&= w_l\theta_1\theta_2 + w_s\theta_1(1-\theta_2) \\
&= (w_l - w_s)\theta_1\theta_2 + w_s\theta_1.
\end{aligned}$$

另一方面, 再次根据双重条件方差公式, 有

$$\begin{aligned}\mathsf{Var}\,(X_{m+1}\,|\,\theta) &= \mathsf{E}[\mathsf{Var}\,(X_{m+1}\,|\,N_{m+1},\theta)\,|\,\theta] + \mathsf{Var}\,[\mathsf{E}(X_{m+1}\,|\,N_{m+1},\theta)\,|\,\theta] \\ &= \mathsf{E}[N_{m+1}\theta_2(1-\theta_2)\,|\,\theta] + \mathsf{Var}\,(N_{m+1}\theta_2\,|\,\theta) \\ &= \theta_1\theta_2(1-\theta_2) + \theta_1\theta_2^2 \\ &= \theta_1\theta_2\end{aligned}$$

以及

$$\begin{aligned}\mathsf{Var}\,(Z_{m+1}\,|\,\theta) &= \mathsf{E}[\mathsf{Var}\,(Z_{m+1}\,|\,N_{m+1},\theta)\,|\,\theta] + \mathsf{Var}\,[\mathsf{E}(Z_{m+1}\,|\,N_{m+1},\theta)\,|\,\theta] \\ &= \mathsf{E}[N_{m+1}\theta_2(1-\theta_2)\,|\,\theta] + \mathsf{Var}\,[N_{m+1}(1-\theta_2)\,|\,\theta] \\ &= \theta_1\theta_2(1-\theta_2) + \theta_1(1-\theta_2)^2 \\ &= \theta_1(1-\theta_2).\end{aligned}$$

根据双重协方差公式得到

$$\begin{aligned}&\mathsf{Cov}\,(X_{m+1},N_{m+1}\,|\,\theta) \\ =\ &\mathsf{E}[\mathsf{Cov}\,(X_{m+1},N_{m+1}\,|\,N_{m+1},\theta)\,|\,\theta] + \mathsf{Cov}\,[\mathsf{E}(X_{m+1}\,|\,N_{m+1},\theta),N_{m+1}\,|\,\theta] \\ =\ &\mathsf{Cov}\,(N_{m+1}\theta_2,N_{m+1}\,|\,\theta) \\ =\ &\theta_1\theta_2.\end{aligned}$$

因此得到

$$\mathsf{Cov}\,(X_{m+1},Z_{m+1}\,|\,\theta) = \mathsf{Cov}\,(X_{m+1},N_{m+1}-X_{m+1}\,|\,\theta) = 0. \tag{9}$$

最后, 经过计算有

$$\begin{aligned}\sigma_v^2(\theta) &= \mathsf{Var}\,(w_lX_{m+1}+w_sZ_{m+1}\,|\,\theta) \\ &= w_l^2\mathsf{Var}\,(X_{m+1}\,|\,\theta) + w_s^2\mathsf{Var}\,(Z_{m+1}\,|\,\theta) + 2w_lw_s\mathsf{Cov}\,(X_{m+1},Z_{m+1}\,|\,\theta) \\ &= w_l^2\theta_1\theta_2 + w_s^2\theta_1(1-\theta_2) \\ &= (w_l^2-w_s^2)\theta_1\theta_2 + w_s^2\theta_1.\end{aligned}$$

引理得证. $\square$

在假设 1–3 中, 分别计算 $\mu_v(\theta)$ 和 $\sigma_v^2(\theta)$ 的后验分布, 可以得到相应的贝叶斯估计, 叙述为下面的定理.

定理 6 在平方损失函数下, 假设 1–3 中参数 $\mu_v(\theta)$ 和 $\sigma_v^2(\theta)$ 的贝叶斯估计为

$$\widehat{\mu_v(\theta)}^{\mathrm{B}} = \frac{(\alpha+N_{\bullet})[(w_la+w_sb)+(w_lX_{\bullet}+w_sZ_{\bullet})]}{(\beta+m)(a+b+N_{\bullet})} \tag{10}$$

以及

$$\widehat{\sigma_v^2(\theta)}^{\mathrm{B}} = \frac{(\alpha+N_\cdot)[w_l^2(a+X_\cdot)+w_s^2(b+Z_\cdot)]}{(\beta+m)(a+b+N_\cdot)}.$$

因此在方差相关保费原理 (6) 中风险保费的贝叶斯估计为

$$\widehat{R_v(\theta)}^{\mathrm{B}} = \widehat{\mu_v(\theta)}^{\mathrm{B}} + g\big(\widehat{\mu_v(\theta)}^{\mathrm{B}}, \widehat{\sigma_v^2(\theta)}^{\mathrm{B}}\big). \tag{11}$$

证明: 在平方损失函数下, 最小化风险函数

$$\min_{h_2(\widetilde{S})} \mathsf{E}\big[(\mu_v(\theta)-h(\widetilde{S}))^2\big] \quad 和 \quad \min_{h_1(\widetilde{S})} \mathsf{E}\big[(\sigma_v^2(\theta)-h_2(\widetilde{S}))^2\big].$$

得到参数 $\mu_v(\theta)$ 和 $\sigma_v^2(\theta)$ 的贝叶斯估计为

$$\widehat{\mu_v(\theta)}^{\mathrm{B}} = \mathsf{E}[\mu_v(\theta)\,|\,\widetilde{S}] \quad 以及 \quad \widehat{\sigma_v^2(\theta)}^{\mathrm{B}} = \mathsf{E}[\sigma_v^2(\theta)\,|\,\widetilde{S}].$$

根据条件分布的链式法则, 参数 (θ_1,θ_2) 的联合后验分布为

$$\begin{aligned}
\pi(\theta_1,\theta_2\,|\,\widetilde{s}) &\propto \pi_1(\theta_1)\pi_2(\theta_2)p(n_1,n_2,\cdots,n_m\,|\,\theta)\\
&\qquad\cdot f(x_{11},\cdots,x_{m,n_m},z_{11},\cdots,z_{m,n_m}\,|\,\theta,n_1,n_2,\cdots,n_m)\\
&\propto \theta_1^{\alpha-1}\mathrm{e}^{-\beta\theta_1}\theta_2^{a-1}(1-\theta_2)^{b-1}\prod_{i=1}^m(\theta_1^{n_i}\mathrm{e}^{-\theta_1})\prod_{i=1}^m\prod_{j=1}^{n_i}[\theta_2^{x_{ij}}(1-\theta_2)^{z_{ij}}]\\
&\propto \theta_1^{\alpha+n_\cdot-1}\mathrm{e}^{-(\beta+m)\theta_1}\theta_2^{a+x_\cdot-1}(1-\theta_2)^{b+z_\cdot-1},
\end{aligned}$$

其中

$$n_\cdot=\sum_{i=1}^m n_i,\quad x_\cdot=\sum_{i=1}^m x_i=\sum_{i=1}^m\sum_{j=1}^{n_i}x_{ij}\quad 以及\quad z_\cdot=\sum_{i=1}^m z_i=\sum_{i=1}^m\sum_{j=1}^{n_i}z_{ij}.$$

容易看出, 在样本给定为观测值 $\widetilde{s}$ 条件下, 参数 θ_1 与 θ_2 后验分布独立, 且有

$$\pi_1(\theta_1\,|\,\widetilde{s}) = \frac{(\beta+m)^{\alpha+n_\cdot}}{\Gamma(\alpha+n_\cdot)}\theta_1^{\alpha+n_\cdot-1}\mathrm{e}^{-(\beta+m)\theta_1}I(\theta_1>0),$$

以及

$$\pi_2(\theta_2\,|\,\widetilde{s}) = \frac{1}{\mathrm{Be}(a+x_\cdot,b+z_\cdot)}\theta_2^{a+x_\cdot-1}(1-\theta_2)^{b+z_\cdot-1}I(0<\theta_2<1).$$

注意到 $x_{ij}+z_{ij}=1$, $x_i+z_i=n_i$ 以及 $x_\cdot+z_\cdot=n_\cdot$, 得到参数 θ_1 与 θ_2 的后验均值估计分别为

$$\widehat{\theta_1}^{\mathrm{B}} = \mathsf{E}(\theta_1\,|\,\widetilde{S}) = \frac{\alpha+N_\cdot}{\beta+m} \quad 以及 \quad \widehat{\theta_2}^{\mathrm{B}} = \mathsf{E}(\theta_2\,|\,\widetilde{S}) = \frac{\alpha+X_\cdot}{a+b+N_\cdot}.$$

则得到

$$\widehat{\mu_v(\theta)}^{\mathrm{B}} = \mathsf{E}\big[(w_l-w_s)\theta_1\theta_2+w_s\theta_1\,|\,\widetilde{S}\big]$$

$$
\begin{aligned}
&= (w_l - w_s)\mathsf{E}(\theta_1 \mid \widetilde{S})\mathsf{E}(\theta_2 \mid \widetilde{S}) + w_s\mathsf{E}(\theta_1 \mid \widetilde{S}) \\
&= (w_l - w_s)\frac{\alpha + N_\cdot}{\beta + m} \times \frac{a + X_\cdot}{a + b + N_\cdot} + w_s\frac{\alpha + N_\cdot}{\beta + m} \\
&= \frac{(\alpha + N_\cdot)[w_l(a + X_\cdot) + w_s(b + Z_\cdot)]}{(\beta + m)(a + b + N_\cdot)} \\
&= \frac{(\alpha + N_\cdot)[(w_l a + w_s b) + (w_l X_\cdot + w_s Z_\cdot)]}{(\beta + m)(a + b + N_\cdot)}
\end{aligned}
$$

以及

$$
\begin{aligned}
\widehat{\sigma_v^2(\theta)}^{\mathrm{B}} &= \mathsf{E}\big[(w_l^2 - w_s^2)\theta_1\theta_2 + w_s^2\theta_1 \mid \widetilde{S}\big] \\
&= (w_l^2 - w_s^2)\frac{\alpha + N_\cdot}{\beta + m} \times \frac{a + X_\cdot}{a + b + N_\cdot} + w_s^2\frac{\alpha + N_\cdot}{\beta + m} \\
&= \frac{(\alpha + N_\cdot)[w_l^2(a + X_\cdot) + w_s^2(b + Z_\cdot)]}{(\beta + m)(a + b + N_\cdot)}.
\end{aligned}
$$

代入后即可得到方差相关保费原理中风险保费的贝叶斯估计. □

注意到参数 $\mu_v(\theta)$ 和 $\sigma_v^2(\theta)$ 关于先验分布的无条件均值分别为

$$
\mu_{v0} = \mathsf{E}[\mu_v(\theta)] = (w_l - w_s)\mathsf{E}(\theta_1)\mathsf{E}(\theta_2) + w_s\mathsf{E}(\theta_1) = \frac{\alpha(w_l a + w_s b)}{\beta(a + b)}
$$

以及

$$
\sigma_{v0}^2 = \mathsf{E}[\sigma_v^2(\theta)] = (w_l^2 - w_s^2)\mathsf{E}(\theta_1)\mathsf{E}(\theta_2) + w_s^2\mathsf{E}(\theta_1) = \frac{\alpha(w_l^2 a + w_s^2 b)}{\beta(a + b)}.
$$

则得到下面的定理.

定理 7 在平方损失函数下, 假设 1–3 中参数 $\mu_v(\theta)$ 和 $\sigma_v^2(\theta)$ 的贝叶斯估计可以表达为下面的信度形式

$$
\widehat{\mu_v(\theta)}^{\mathrm{B}} = Z_v h_1(N_\cdot, X_\cdot, Z_\cdot) + (1 - Z_v)\mu_{v0} \tag{12}
$$

以及

$$
\widehat{\sigma_v^2(\theta)}^{\mathrm{B}} = Z_v h_2(N_\cdot, X_\cdot, Z_\cdot) + (1 - Z_v)\sigma_{v0}^2. \tag{13}
$$

其中

$$
\begin{aligned}
h_1(N_\cdot, X_\cdot, Z_\cdot) &= \frac{(\alpha + N_\cdot)(w_l X_\cdot + w_s Z_\cdot) + N_\cdot(w_l a + w_s b)}{m(a + b + N_\cdot) + \beta N_\cdot}, \\
h_2(N_\cdot, X_\cdot, Z_\cdot) &= \frac{(\alpha + N_\cdot)(w_l^2 X_\cdot + w_s^2 Z_\cdot) + N_\cdot(w_l^2 a + w_s^2 b)}{m(a + b + N_\cdot) + \beta N_\cdot}
\end{aligned}
$$

以及

$$
Z_v = \frac{m(a + b + N_\cdot) + \beta N_\cdot}{m(a + b + N_\cdot) + \beta N_\cdot + \beta(a + b)}.
$$

证明: 根据定理 6, 有

$$\begin{aligned}\widehat{\mu_v(\theta)}^{\mathrm{B}} &= \frac{(\alpha+N_{\boldsymbol{\cdot}})[(w_la+w_sb)+(w_lX_{\boldsymbol{\cdot}}+w_sZ_{\boldsymbol{\cdot}})]}{(\beta+m)(a+b+N_{\boldsymbol{\cdot}})}\\ &= \frac{\alpha(w_la+w_sb)}{m(a+b+N_{\boldsymbol{\cdot}})+\beta N_{\boldsymbol{\cdot}}+\beta(a+b)}+\frac{(\alpha+N_{\boldsymbol{\cdot}})(w_lX_{\boldsymbol{\cdot}}+w_sZ_{\boldsymbol{\cdot}})+N_{\boldsymbol{\cdot}}(w_la+w_sb)}{m(a+b+N_{\boldsymbol{\cdot}})+\beta N_{\boldsymbol{\cdot}}+\beta(a+b)}\\ &= (1-Z_v)\mu_{v0}+Z_vh_1(N_{\boldsymbol{\cdot}},X_{\boldsymbol{\cdot}},Z_{\boldsymbol{\cdot}}).\end{aligned}$$

类似地可以证明式 (13). □

注记 8 在对风险保费的估计过程中, 若无任何样本信息, 则 μ_{v0} 和 σ_{v0}^2 分别是 $\mu_{v0}(\theta)$ 和 $\sigma_v^2(\theta)$ 的聚合估计, 则称 $R_0=\mu_{v0}+g(\mu_{v0},\sigma_{v0}^2)$ 为聚合保费. 而 $h_1(N_{\boldsymbol{\cdot}},X_{\boldsymbol{\cdot}},Z_{\boldsymbol{\cdot}})$ 与 $h_2(N_{\boldsymbol{\cdot}},X_{\boldsymbol{\cdot}},Z_{\boldsymbol{\cdot}})$ 是 $\mu_{v0}(\theta)$ 和 $\sigma_v^2(\theta)$ 经验估计. 定理 7 说明, 参数 $\mu_{v0}(\theta)$ 和 $\sigma_v^2(\theta)$ 的贝叶斯估计均能表达为聚合估计和经验估计的加权平均, 其中权重 $0\leqslant Z_v\leqslant 1$, 且 Z_v 是样本容量 m 的增函数.

注记 9 在传统的信度理论中, Bühlmann[19] 给出了净保费原理下风险保费的信度估计, 表达为

$$\widehat{\mu(\theta)}=Z\overline{X}+(1-Z)\mu_0, \tag{14}$$

其中 $\mu(\theta)=\mathsf{E}(X\,|\,\theta)$, $\mu_0=\mathsf{E}[\mu(\theta)]$, $\overline{X}=m^{-1}\sum\limits_{i=1}^{m}X_i$, 而

$$Z=\frac{m\mathsf{Var}\,[\mu(\theta)]}{m\mathsf{Var}\,[\mu(\theta)]+\mathsf{E}[\mathsf{Var}\,(X\,|\,\theta)]}$$

为信度因子. 注意到, 本模型中, 不仅将经典的信度理论推广到聚合风险模型, 而且将净保费原理推广到方差相关保费原理. 这时得到的保费估计仍然能表达为经验保费和聚合保费的加权平均. 但传统信度因子不依赖于样本观测值, 是非随机变量. 而定理 7 中的权重因子 Z_v 依赖于索赔样本 $N_{\boldsymbol{\cdot}}$, 因此是随机变量. 然而, 我们仍然可以验证: 当 $m\to 0$ 时 $Z_v\to 0$, a.s.; 当 $m\to\infty$ 时 $Z_v\to 1$, a.s.. 这与传统的信度因子解释相吻合: 样本容量越大, 则在经验估计上分配较大的权重, 当样本容量较小时, 在聚合估计上分配更多的权重. 因而也称 Z_v 为信度因子.

定理 10 参数 $\mu_v(\theta)$ 和 $\sigma_v^2(\theta)$ 的贝叶斯估计都是强相合的, 即当 θ 给定且 $m\to\infty$ 有

$$\widehat{\mu_v(\theta)}^{\mathrm{B}}\xrightarrow{\text{a.s.}}\mu_v(\theta),\qquad \widehat{\sigma_v^2(\theta)}^{\mathrm{B}}\xrightarrow{\text{a.s.}}\sigma_v^2(\theta).$$

进而, 若二元函数 $g(x,y)$ 是连续函数, 则有 $\widehat{R_v(\theta)}^{\mathrm{B}}\xrightarrow{\text{a.s.}}R_v(\theta)$.

证明: 由于当 θ 给定 (N_i,X_i,Z_i) 对 $i=1,2,\cdots,m$ 相互独立且同分布, 其中 $X_i=\sum\limits_{j=1}^{N_i}X_{ij}$ 以及 $Z_i=\sum\limits_{j=1}^{N_i}Z_{ij}$. 注意到 $\mathsf{E}(N_i\,|\,\theta)=\theta_1$, 则有

$$\mathsf{E}(X_i\,|\,\theta)=\mathsf{E}[(X_i\,|\,N_i,\theta)\,|\,\theta]=\theta_1\theta_2$$

以及

$$\mathsf{E}(Z_i\,|\,\theta)=\mathsf{E}[(Z_i\,|\,N_i,\theta)\,|\,\theta]=\theta_1(1-\theta_2),$$

其中 $\overline{X}=m^{-1}\sum\limits_{i=1}^{m}X_i$, $\overline{Z}=m^{-1}\sum\limits_{i=1}^{m}Z_i$, $\overline{N}=m^{-1}\sum\limits_{i=1}^{m}N_i$, 则根据强大数定律, 有

$$\overline{N}\to\theta_1,\quad \overline{X}\to\theta_1\theta_2,\quad \overline{Z}\to\theta_1(1-\theta_2),\qquad \text{a.s..}$$

因此

$$\begin{aligned}h_1(N_\cdot,X_\cdot,Z_\cdot)&=\frac{(\alpha/m+\overline{N})(w_l\overline{X}+w_s\overline{Z})+(w_la+w_sb)\overline{N}/m}{[(a+b)/m+\overline{N}]+\beta\overline{N}/m}\\&\to w_l\theta_1\theta_2+w_s\theta_1(1-\theta_2)\\&=\mu_v(\theta),\qquad \text{a.s.}\end{aligned}$$

以及

$$\begin{aligned}h_2(N_\cdot,X_\cdot,Z_\cdot)&=\frac{(\alpha/m+\overline{N})(w_l^2\overline{X}+w_s^2\overline{Z})+\overline{N}(w_l^2a+w_s^2b)}{[(a+b)/m+\overline{N}]+\beta\overline{N}/m}\\&\to w_l^2\theta_1\theta_2+w_s^2\theta_1(1-\theta_2)\\&=\sigma_v^2(\theta),\qquad \text{a.s..}\end{aligned}$$

类似地, 容易验证

$$Z_v=\frac{m(a+b+N_\cdot)+\beta N_\cdot}{m(a+b+N_\cdot)+\beta N_\cdot+\beta(a+b)}=\frac{(a+b)/m+N_\cdot/m+\beta N_\cdot/m^2}{(a+b)/m+N_\cdot/m+\beta(a+b)/m^2}\to 1,\quad \text{a.s..}$$

则根据几乎处处收敛的性质, 有

$$\widehat{\mu_v(\theta)}^{\mathrm{B}}=Z_vh_1(N_\cdot,X_\cdot,Z_\cdot)+(1-Z_v)\mu_{v0}\to\mu_v(\theta),\qquad \text{a.s.}$$

以及

$$\widehat{\sigma_v^2(\theta)}^{\mathrm{B}}\xrightarrow{\text{a.s.}}\sigma_v^2(\theta).$$

再根据连续性定理可得 $\widehat{R_v(\theta)}^{\mathrm{B}}\xrightarrow{\text{a.s.}}R_v(\theta)$. □

根据中心极限定理和 Crámer 定理 [20], 可以证明下面的渐近正态性.

定理 11　参数 $\mu_v(\theta)$ 和 $\sigma_v^2(\theta)$ 的贝叶斯估计都是渐近正态的, 即有

$$\frac{\sqrt{m}\big[\widehat{\mu_v(\theta)}^{\mathrm{B}}-\mu_v(\theta)\big]}{\sigma_v(\theta)}\xrightarrow{\mathrm{L}}\mathrm{N}(0,1)\tag{15}$$

以及

$$\frac{\sqrt{m}\big[\widehat{\sigma_v^2(\theta)}^{\mathrm{B}}-\sigma_v^2(\theta)\big]}{\sigma_{vv}(\theta)}\xrightarrow{\mathrm{L}}\mathrm{N}(0,1),\tag{16}$$

其中 $\sigma_{vv}^2(\theta)=w_l^4\theta_1\theta_2+w_s^4\theta_1(1-\theta_2)$. 进而, 有

$$\frac{\sqrt{m}\big[\widehat{R_v(\theta)}^{\mathrm{B}}-R_v(\theta)\big]}{\sigma_\psi(\theta)}\xrightarrow{\mathrm{L}}\mathrm{N}(0,1),\tag{17}$$

其中

$$\begin{aligned}\sigma_\psi^2(\theta)=&[1+\dot{g}_x(\theta)]^2[w_l^2\theta_1\theta_2+w_s^2\theta_1(1-\theta_2)]\\&+2\dot{g}_y(\theta)[1+\dot{g}_x(\theta)][w_l^3\theta_1\theta_2+w_s^3\theta_1(1-\theta_2)]+(\dot{g}_y)^2[w_l^4\theta_1\theta_2+w_s^4\theta_1(1-\theta_2)].\end{aligned}$$

证明: 记 $T_i=w_lX_i+w_sZ_i$, $i=1,2,\cdots,m$, 则有

$$\overline{T}=\frac{1}{m}\sum_{i=1}^m T_i=w_l\overline{X}+w_s\overline{Z}.$$

在 θ 给定条件下 $T_1,T_2,\cdots,T_m$ 相互独立且具有共同的分布. 注意到在 θ 给定条件下

$$\mathsf{Cov}\,(X_i,Z_i\,|\,\theta)=0,$$

因此可以得到 T_i 的条件期望和条件方差为

$$\mathsf{E}(T_i\,|\,\theta)=w_l\mathsf{E}(X_i\,|\,\theta)+w_s\mathsf{E}(Z_i\,|\,\theta)=w_l\theta_1\theta_2+w_s\theta_1(1-\theta_2)=\mu_v(\theta)$$

以及

$$\mathsf{Var}\,(T_i\,|\,\theta)=w_l^2\mathsf{Var}\,(X_i\,|\,\theta)+w_s^2\mathsf{Var}\,(Z_i\,|\,\theta)=w_l^2\theta_1\theta_2+w_s^2\theta_1(1-\theta_2)=\sigma_v^2(\theta).$$

根据中心极限定理, 有

$$\frac{\sqrt{m}\,[\overline{T}-\mu_v(\theta)]}{\sigma_v(\theta)}\xrightarrow{\mathrm{L}}\mathrm{N}(0,1).$$

又因为

$$\begin{aligned}&\sqrt{m}\big[\widehat{\mu_v(\theta)}^{\mathrm{B}}-\mu_v(\theta)\big]-\sqrt{m}\Big[\frac{N_\cdot(w_lX_\cdot+w_sZ_\cdot)}{(\beta+m)(a+b+N_\cdot)}-\mu_v(\theta)\Big]\\=&\sqrt{m}\Big\{\frac{(\alpha+N_\cdot)[(w_la+w_sb)+(w_lX_\cdot+w_sZ_\cdot)]}{(\beta+m)(a+b+N_\cdot)}-\frac{N_\cdot(w_lX_\cdot+w_sZ_\cdot)}{(\beta+m)(a+b+N_\cdot)}\Big\}\\=&\sqrt{m}\,\frac{(w_la+w_sb)(\alpha+\overline{N})+\alpha(w_l\overline{X}+w_s\overline{Z})}{(\beta+m)[(a+b)/m+\overline{N}]}\\\to&\,0\end{aligned}$$

以及

$$\sqrt{m}\Big[\frac{N_\cdot(w_lX_\cdot+w_sZ_\cdot)}{(\beta+m)(a+b+N_\cdot)}-\mu_v(\theta)\Big]-\sqrt{m}\,[\overline{T}-\mu_v(\theta)]$$

$$
\begin{aligned}
&= \frac{N_{\cdot}(w_l X_{\cdot} + w_s Z_{\cdot})}{(\beta+m)(a+b+N_{\cdot})} - (w_l \overline{X} + w_s \overline{Z}) \\
&= \sqrt{m}\,\frac{mN_{\cdot} - (\beta+m)(a+b+N_{\cdot})}{(\beta+m)(a+b+N_{\cdot})}(w_l \overline{X} + w_s \overline{Z}) \\
&= \sqrt{m}\,\frac{-(\beta+m)(a+b) + \beta N_{\cdot}}{(\beta+m)(a+b+N_{\cdot})}(w_l \overline{X} + w_s \overline{Z}) \\
&= \frac{1}{\sqrt{m}}\frac{-(\beta/m+1)(a+b) + \beta \overline{N}}{(\beta/m+1)[(a+b)/m + \overline{N}]}(w_l \overline{X} + w_s \overline{Z}) \\
&\to 0.
\end{aligned}
$$

因此, 根据 Slustky 定理 [20], 可得 (15) 式. 类似地, 记

$$U_i = w_l^2 X_i + w_s^2 Z_i, \qquad i = 1, 2, \cdots, m,$$

则 $U_1, U_2, \cdots, U_m$ 条件独立同分布且

$$\mathsf{E}(U_i \,|\, \theta) = \sigma_v^2(\theta) \quad \text{以及} \quad \mathsf{Var}\,(U_i \,|\, \theta) = w_l^4 \theta_1 \theta_2 + w_s^4 \theta_1 (1-\theta_2) = \sigma_{vv}^2(\theta).$$

则类似地可证明 (16) 式. 进一步地, 记

$$Q_i = (T_i, U_i)^{\mathsf{T}} = (w_l X_i + w_s Z_i, w_l^2 X_i + w_s^2 Z_i), \qquad i = 1, 2, \cdots, m.$$

则在 θ 给定条件下 $Q_1, Q_2, \cdots, Q_m$ 条件独立且具有共同的条件分布, 容易验证

$$\mathsf{E}(Q_i \,|\, \theta) = (\mu_v(\theta), \sigma_v^2(\theta))^{\mathsf{T}}$$

以及

$$\mathsf{Var}\,(Q_i \,|\, \theta) = \begin{pmatrix} w_l^2 \theta_1 \theta_2 + w_s^2 \theta_1 (1-\theta_2) & w_l^3 \theta_1 \theta_2 + w_s^3 \theta_1 (1-\theta_2) \\ w_l^3 \theta_1 \theta_2 + w_s^3 \theta_1 (1-\theta_2) & w_l^4 \theta_1 \theta_2 + w_s^4 \theta_1 (1-\theta_2) \end{pmatrix} \triangleq \Phi(\theta).$$

根据随机向量的中心极限定理, 有

$$\sqrt{m}\,[\overline{Q} - \mathsf{E}(Q_1 \,|\, \theta)] \xrightarrow{\mathrm{L}} \mathrm{N}(0, \Phi(\theta)).$$

即

$$\sqrt{m}\left[\begin{pmatrix} w_l \overline{X} + w_s \overline{Z} \\ w_l^2 \overline{X} + w_s^2 \overline{Z} \end{pmatrix} - \begin{pmatrix} \mu_v(\theta) \\ \sigma_v^2(\theta) \end{pmatrix}\right] \xrightarrow{\mathrm{L}} \mathrm{N}(0, \Phi(\theta)).$$

根据前面的计算, 可以证明

$$\sqrt{m}\left[\begin{pmatrix} w_l \overline{X} + w_s \overline{Z} \\ w_l^2 \overline{X} + w_s^2 \overline{Z} \end{pmatrix} - \begin{pmatrix} \widehat{\mu_v(\theta)}^{\mathrm{B}} \\ \widehat{\sigma_v^2(\theta)}^{\mathrm{B}} \end{pmatrix}\right] \xrightarrow{\mathrm{L}} 0.$$

再次根据 Slustky 定理, 得到

$$\sqrt{m}\left[\begin{pmatrix}\widehat{\mu_v(\theta)}^{\mathrm{B}}\\ \widehat{\sigma_v^2(\theta)}^{\mathrm{B}}\end{pmatrix}-\begin{pmatrix}\mu_v(\theta)\\ \sigma_v^2(\theta)\end{pmatrix}\right]\xrightarrow{\mathrm{L}}\mathrm{N}(0,\Phi(\theta)).$$

令 $\varphi(x,y)=x+g(x,y)$, 则

$$\frac{\partial\varphi}{\partial x}=1+\dot{g}_x(x,y),\qquad \frac{\partial\varphi}{\partial y}=\dot{g}_y(x,y).$$

记

$$\dot{g}_x(\theta)=\dot{g}_x(\mu_v(\theta),\sigma_v^2(\theta)),\qquad \dot{g}_y(\theta)=\dot{g}_y(\mu_v(\theta),\sigma_v^2(\theta)),$$

则 $\varphi(x,y)$ 的偏导数在点 $(\mu_v(\theta),\sigma_v^2(\theta))$ 处的值为:

$$\frac{\partial\varphi}{\partial x}(\mu_v(\theta),\sigma_v^2(\theta))=1+\dot{g}_x(\theta)\quad 以及\quad \frac{\partial\varphi}{\partial y}(\mu_v(\theta),\sigma_v^2(\theta))=\dot{g}_y(\theta).$$

注意到

$$\begin{aligned}&(1+\dot{g}_x(\theta),\dot{g}_y(\theta))\Phi(\theta)\begin{pmatrix}1+\dot{g}_x(\theta)\\ \dot{g}_y(\theta)\end{pmatrix}\\ =&[1+\dot{g}_x(\theta)]^2[w_l^2\theta_1\theta_2+w_s^2\theta_1(1-\theta_2)]+2\dot{g}_y(\theta)[1+\dot{g}_x(\theta)][w_l^3\theta_1\theta_2+w_s^3\theta_1(1-\theta_2)]\\ &+(\dot{g}_y)^2[w_l^4\theta_1\theta_2+w_s^4\theta_1(1-\theta_2)]\\ =&\sigma_\psi^2(\theta),\end{aligned}$$

根据随机向量的 Crámer 定理 [20] 可得

$$\sqrt{m}\big[\varphi\big(\widehat{\mu_v(\theta)}^{\mathrm{B}},\widehat{\sigma_v^2(\theta)}^{\mathrm{B}}\big)-\varphi\big(\mu_v(\theta),\sigma_v^2(\theta)\big)\big]\xrightarrow{\mathrm{L}}\mathrm{N}\left(0,(1+\dot{g}_x(\theta),\dot{g}_y(\theta))\Phi(\theta)\begin{pmatrix}1+\dot{g}_x(\theta)\\ \dot{g}_y(\theta)\end{pmatrix}\right),$$

即证明了 (17) 式. □

§4. 实证分析和模拟

在实际运用中, 第三节得到的贝叶斯估计中仍然存在一些未知参数, 这些参数是贝叶斯假设中先验分布的超参数, 在非寿险精算中称为结构参数. 为了估计这些结构参数, 需要根据一定的统计方法利用实际数据对模型进行拟合.

本数据来源于文献 [21], 可从网站 http://blog.sina.com.cn/mengshw 进行下载. 数据集包含 dat2 和 dat3 两个数据集, 其中 dat2 是 413 169 张汽车第三者责任保险保单的的索赔次数数据, dat3 记录了数据集 dat2 中汽车第三者责任保险的所有保单每次的索赔金额. 根据这两个数据, 可以得到下面的分析结果.

表 1　索赔次数 $\{N_1, N_2, \cdots, N_m\}$ 的观测值统计

k	0	1	2	3	4
n_k	397 779	14 633	726	28	3
f_k	96.27513	7.541615	0.17571	0.00678	0.00073

其中 n_k 和 f_k 分别表示索赔次数样本 $\{N_1, N_2, \cdots, N_m\}$ 中取值为 k 的样本个数以及相应的频率. 根据上表一的数据可以计算得到索赔次数 N 的样本均值和样本方差分别为

$$\overline{N} = \sum_{k=1}^{4} k f_k = 0.03916 \quad \text{以及} \quad \widehat{V} = \sum_{k=1}^{4} k^2 f_k - \overline{N}^2 = 0.04164.$$

注意到在假设 1–3 成立条件下有

$$\mathsf{E}N = \mathsf{E}[\mathsf{E}(N \,|\, \theta)] = \mathsf{E}(\theta_1) = \frac{\alpha}{\beta}$$

以及

$$\mathsf{Var}\,(N) = \mathsf{E}[\mathsf{Var}\,(N \,|\, \theta)] + \mathsf{Var}\,[\mathsf{E}(N \,|\, \theta)] = \mathsf{E}(\theta_1) = \frac{\alpha(1+\beta)}{\beta^2}.$$

令 $\mathsf{E}N = \overline{N}$, $\mathsf{Var}\,(N) = \widehat{V}$, 解得

$$\widehat{\alpha} = \frac{\overline{N}^2}{\widehat{V} - \overline{N}} = 0.6199 \quad \text{以及} \quad \widehat{\beta} = \frac{\overline{N}}{\widehat{V} - \overline{N}} = 15.83.$$

如果取 $d = 1\,000$, 即超过索赔额 1 000 即认为是大额索赔, 否则认为是小额索赔, 对 data2 中的数据进行统计, 得到 $X_1, X_2, \cdots, X_m$ 的频率表.

表 2　大额索赔次数 $\{X_1, X_2, \cdots, X_m\}$ 的频率表

k	0	1	2	3
r_k	412 627	525	15	2
p_k	99.8688	0.1271	0.0036	0.0005

在表 2 中, r_k 和 p_k 分别表示大额索赔样本 $\{X_1, X_2, \cdots, X_m\}$ 中取值为 k 的样本个数以及相应的频率. 计算得到 $\{X_1, X_2, \cdots, X_m\}$ 的样本均值和样本方差分别为

$$\overline{X} = \frac{1}{m} \sum_{i=1}^{m} X_i = \sum_{k=1}^{3} k f_k = 0.001358 \quad \text{和} \quad \widehat{W} = \sum_{k=1}^{3} k^2 r_k - \overline{X}^2 = 0.001548.$$

根据假设 1–3, 有

$$\mathsf{E}X_1 = \mathsf{E}[\mathsf{E}(X_1 \,|\, \theta)] = \mathsf{E}(\theta_1 \theta_2) = \frac{\alpha a}{\beta(a+b)}$$

以及

$$\begin{aligned}\mathsf{Var}\,(X_1) &= \mathsf{E}[\mathsf{Var}\,(X_1\,|\,\theta)] + \mathsf{Var}\,[\mathsf{E}(X_1\,|\,\theta)]\\ &= \mathsf{E}(\theta_1\theta_2) + \mathsf{Var}\,(\theta_1\theta_2)\\ &= \frac{\alpha a}{\beta(a+b)} + \frac{\alpha ab}{\beta^2(a+b)^2(a+b+1)} + \frac{\alpha a^2}{\beta^2(a+b)^2} + \frac{\alpha^2 ab}{\beta^2(a+b)^2(a+b+1)}\\ &= \frac{\alpha a[\alpha(a+b+1)+\beta(a+b)(a+b+1)+b(\alpha+1)]}{\beta^2(a+b)^2(a+b+1)}.\end{aligned}$$

根据矩估计思想, 令 $\mathsf{E}X_1 = \overline{X}$, 则得到 $b = a\cdot c_1$, 其中 $c_1 = (\alpha - \overline{X}\beta)/(\overline{X}\beta)$. 再令 $\mathsf{Var}\,(X_1) = \widehat{W}$, 则得到

$$a = \frac{\alpha c_1(1+\alpha) - \beta^2(1+c_1)^2 c_2}{\beta^2(1+c_1)^3 c_2}$$

以及

$$c_1 = \frac{\alpha - \overline{X}\beta}{\overline{X}\beta}, \qquad c_2 = \widehat{M} - \frac{\alpha}{\beta(1+c_1)} - \frac{\alpha}{\beta^2(1+c_1)^2}.$$

将 $\widehat{\alpha}$、$\widehat{\beta}$ 代入后得到 $\widehat{a} = 0.02485$ 以及 $\widehat{b} = 0.6919$.

根据真实数据得到的 $\widehat{\alpha}$、$\widehat{\beta}$、$\widehat{a}$、$\widehat{b}$, 下面对四种保费原理中风险保费的贝叶斯估计进行数值模拟, 验证相应的统计性质. 取 $w_l = 1.5$、$w_s = 0.8$, 对不同的二元函数 $g_1(x,y) = \alpha_1 x$, $g_2(x,y) = \alpha_2 y$, $g_3(x,y) = \alpha_3\sqrt{y}$ 以及 $g_4(x,y) = \alpha_4 y/x$, 分别对应期望值保费原理、方差保费原理、标准差保费原理以及修正方差保费原理. 取不同的 θ 分别计算贝叶斯估计条件均方标准误差, 定义为

$$\mathrm{std}_k(\theta) = \sqrt{\mathsf{E}\big[\big(\widehat{R_v(\theta)}^{\mathrm{B}} - R_v(\theta)\big)^2\,\big|\,\theta\big]},$$

其中 $k = 1,2,3,4$ 分别表示期望值保费原理、方差保费原理、标准差保费原理和修正方差保费原理中贝叶斯估计的条件均方误差. 为了方便计算和比较, 在四种保费原理中, 取安全负荷系数均 $\alpha_k = 0.5$, $k = 1,2,\cdots,4$. 每次模拟重复 10 000 次, 得到四种方差保费原理下风险保费的贝叶斯估计条件均方误差如下表.

表 3　风险保费的贝叶斯估计及其均方标准差

$\mathrm{std}(\theta)$	$m=10$			$m=50$			$m=200$		
(θ_1,θ_2)	$(3,0.1)$	$(5,0.3)$	$(8,0.8)$	$(3,0.1)$	$(5,0.3)$	$(8,0.8)$	$(3,0.1)$	$(5,0.3)$	$(8,0.8)$
$\mathrm{std}_1(\theta)$	2.362	7.629	9.999	0.918	1.843	7.959	0.272	0.592	1.255
$\mathrm{std}_2(\theta)$	2.303	7.807	11.398	0.898	1.917	7.514	0.270	0.619	1.432
$\mathrm{std}_3(\theta)$	1.866	7.537	7.411	0.711	1.385	2.896	0.211	0.443	0.914
$\mathrm{std}_4(\theta)$	1.579	7.090	6.668	0.614	1.230	2.640	0.183	0.396	0.838

从表 3 的结果看, 当参数 θ 给定时, 四种保费原理中贝叶斯估计的条件均方误差均随着样本的增大而显著较小, 说明这贝叶斯估计是风险保费的相合估计.

参 考 文 献

[1] LEMAIRE J. Bonus-malus systems [M] // TEUGELS J L, SUNDT B. (eds.) *Encyclopedia of Actuarial Science*. New York: Wiley, 2004: 127–139.

[2] FRANGOS N E, VRONTOS S D. Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance [J]. *Astin Bull*, 2001, **31(1)**: 1–22.

[3] GÓMEZ-DÉNIZ E, HERNÁNDEZ-BASTIDA A, FERNÁNDEZ-SÁNCHEZ M P. Computing credibility bonus-malus premiums using the total claim amount distribution [J]. *Hacet J Math Stat*, 2014, **43(6)**: 1047–1061.

[4] GÓMEZ-DÉNIZ E. Bivariate credibility bonus-malus premiums distinguishing between two types of claims [J]. *Insurance Math Econom*, 2016, **70**: 117–124.

[5] LIU H Y, WANG R D. Collective risk models with dependence uncertainty [J]. *Astin Bull*, 2017, **47(2)**: 361–389.

[6] COSSETTE H, MARCEAU E, MTALAI I. Collective risk models with dependence [J]. *Insurance Math Econom*, 2019, **87**: 153–168.

[7] 章溢, 周金亮. 索赔次数的贝叶斯预测与信度近似 [J]. 江西师范大学学报 (自然科学版), 2021, **45(4)**: 353–361.

[8] HASHORVA E, RATOVOMIRIJA G, TAMRAZ M. On some new dependence models derived from multivariate collective models in insurance applications [J]. *Scand Actuar J*, 2017, **2017(8)**: 730–750.

[9] 张良超, 周金亮, 温利民, 零膨胀泊松模型中风险参数的贝叶斯估计 [J]. 江西师范大学学报 (自然科学版), 2020, **44(3)**: 269–274.

[10] YOUNG V R. Premium principles [M] // TEUGELS J L, SUNDT B. (eds.) *Encyclopedia of Actuarial Science*. New York: Wiley, 2004: 1322–1331.

[11] BÜHLMANN H. *Mathematical Methods in Risk Theory* [M]. New York: Springer-Verlag, 1970.

[12] GUERRA M, CENTENO M L. Optimal reinsurance for variance related premium calculation principles [J]. *Astin Bull*, 2010, **40(1)**: 97–121.

[13] CHI Y C. Optimal reinsurance under variance related premium principles [J]. *Insurance Math Econom*, 2012, **51(2)**: 310–321.

[14] 余君, 温利民. 方差相关原理下相依聚合风险模型的贝叶斯保费 [J]. 华东师范大学学报 (自然科学版), 2014, **(4)**: 26–38.

[15] 卢仿先, 曾庆五. 寿险精算数学 [M]. 天津: 南开大学出版社, 2001.

[16] LIANG X Q, LIANG Z B, YOUNG V R. Optimal reinsurance under the mean-variance premium principle to minimize the probability of ruin [J]. *Insurance Math Econom*, 2020, **92**: 128–146.

[17] SUN Z Y, ZHENG X X, ZHANG X. Robust optimal investment and reinsurance of an insurer under variance premium principle and default risk [J]. *J Math Anal Appl*, 2017, **446(2)**: 1666–1686.

[18] 温利民, 李俊雪, 王正武, 等. 在帕累托模型中风险度量的统计分析 [J]. 江西师范大学学报 (自然科学版), 2021, **45(2)**: 211–216.

[19] BÜHLMANN H. Experience Rating and Credibility [J]. *Astin Bull*, 1967, **4(3)**: 199–207.
[20] 茆诗松, 王静龙, 濮晓龙. 高等数理统计 [M]. 北京: 高等教育出版社, 1998.
[21] 孟生旺. 回归模型 [M]. 北京: 中国人民大学出版社, 2015.

The Posterior Ratemaking of Premium in Binary Bayesian Collective Risk Model

ZHANG Yi
(*School of Finance, Jiangxi Normal University, Nanchang, 330022, China;*
School of Finance, Jiangxi University of Finance and Economics, Nanchang, 330013, China)

Abstract: In the Collective risk model, the claim amount is divided into large claims and small claims. Under the variance-related premium principle, the Bayesian estimation of the risk premium in the binary Bayesian collective risk model is derived. The conclusion shows that both the conditional expectation and conditional variance parts of risk premium can be expressed as a weighted form of sample function and aggregate premium, where the weight satisfies the property of "credibility factor". Furthermore, the strong consistency and asymptotic normality of Bayesian estimation is proved. Finally, the method of numerical simulation is used to verify the large sample properties of Bayesian estimation.
Keywords: collective risk model; risk premium; variance-related premium principle; credibility estimation; asymptotic normality
2020 Mathematics Subject Classification: 62G35