文章编号:1000-5862(2018)04-0361-05

索赔次数的加权信度估计及其统计分析

章　溢,李志龙*

(江西财经大学统计学院,江西 南昌　330013)

摘要:利用信度估计的思想研究了索赔次数的信度估计问题,在加权求和损失函数下,最小化期望损失函数得到了索赔次数的概率分布的信度估计. 进而,在索赔次数的 3 大离散分布中分别讨论了风险参数的信度估计,得到了结构参数的估计,并证明了估计的无偏性. 最后,利用数值模拟的方法将本文的信度估计与经典的信度估计进行了比较,验证了本文给出的估计的收敛性质.

关键词:索赔次数;风险参数;信度估计;结构参数

中图分类号:O 211.9　　文献标志码:A　　DOI:10.16357/j.cnki.issn1000-5862.2018.04.07

0　引言

在非寿险精算中,保险公司常关心某个保单的索赔次数. 由于保单的索赔次数反映了该保单的风险大小. 在对非寿险保险保单进行保费厘定时,常常要考虑同类风险的平均索赔情况,并结合该保单在过去若干年的索赔次数的样本记录对保费进行调整. 这种保费定价的方法常称为经验厘定或奖惩系统[1]. 由于风险的非齐次性[2-3],保单的索赔次数 N 常依赖于某个随机参数 θ. 这个随机参数 θ 被称为风险参数,具有某个先验分布 $\pi(\theta)$. 记 $f(x,\theta) = P(N = x \mid \theta)$ 表示索赔次数的条件概率分布. 一般地,风险参数 θ 反映了该保单的风险状况,如在汽车保险中,风险参数表示影响该汽车索赔的所有风险因素(如汽车的型号、性能、里程数或驾驶人的性别、习惯等的综合). 该保单的价格常常是风险参数 θ 的某个函数 $g(\theta)$. 由于风险参数 θ 一般是不可观测的,因而 $g(\theta)$ 需要根据已有的信息进行估计.

在对 $g(\theta)$ 估计的过程中,有 2 类信息可以使用. 一类是风险参数 θ 的先验信息 $\pi(\theta)$,另一类是该保单的索赔次数信息 $N_1,N_2,\cdots,N_n$. 这里 N_i 表示第 i 年的索赔次数,n 为观测到索赔次数的年份. 这种同时结合先验信息和样本信息得到保费的厘定的方法称为信度理论[3]. 信度理论起源于文献[4-5],经过 1 个世纪的发展至今已经是保费定价中的最重要方法[6-7].

然而,大部分信度理论都建立在某些保费原理之上,文献[1]建立了净保费原理的信度模型,文献[8-9]研究了 Esscher 保费原理中的信度估计,文献[10-11]研究了指数保费原理中的信度估计,文献[12]讨论了零效用保费原理的贝叶斯保费等. 注意到不论是什么保费原理,其风险保费总依赖于某个风险参数 θ,而 θ 又是分布 $f(x,\theta)$ 的某个泛函. 因此待估参数总能表达为分布 $f(x,\theta)$ 的某个泛函,不妨记待估参数为 $\eta = h(f(x,\theta))$. 因此只要能得到 $f(x,\theta)$ 的某种“信度”估计,则可以得到 η 的估计 $\hat{\eta} = h(\hat{f}(x,\theta))$. 本文将利用信度理论的思想,建立索赔次数中风险参数的贝叶斯模型,利用加权损失函数得到 $f(x,\theta)$ 的加权信度估计,从而得到待估参数的信度估计,并证明估计的统计性质.

1　索赔次数分布的加权信度估计

假设有 K 个相互独立的保单风险,第 i 份保单在第 j 年的索赔次数为 N_{ij},$j = 1,2,\cdots,n$. 这里 N_{ij} 为取非负整数值的随机变量,假设其概率分布为

$$P(N_{ij} = k) = f(x,\theta_i),x = 1,2,\cdots.$$

由于风险的非齐次性,假设 θ_i,$i = 1,2,\cdots,K$ 为相互独立的随机变量,具有共同的先验分布 $\pi(\theta)$.

收稿日期:2018-04-23

基金项目:国家自然科学基金(71761019)和江西省自然科学基金(20171ACB21022)资助项目.

通信作者:李志龙(1977-),男,江西新干人,教授,博士,博士生导师,主要从事随机分析、数理金融和非线性泛函分析等方面的研究. E-mail:lzl771218@sina.com

本文的目标是根据样本数据$\{N_{ij},i=1,2,\cdots,K,j=1,2,\cdots,n\}$估计$\theta$的某个函数$g(\theta_1,\theta_2,\cdots,\theta_K)$. 由于$\theta_i$一般是其分布$f(x,\theta_i)$的某个泛函,记为$\theta_i=R(f(x,\theta_i))$. 因此将利用信度理论的思想估计$f(x,\theta_i)$,从而得到$g(\theta_1,\theta_2,\cdots,\theta_K)$的估计.

注意到$f(x,\theta_i)=E[I(N_{ij}=x\mid\theta_i)]$,记$f_0(x)=E[f(x,\theta_i)]$,$\tau^2(x)=\mathrm{Var}[f(x,\theta_i)]$,$\sigma^2(x)=E[f(x,\theta_i)(1-f(x,\theta_i))]$. 为估计$f(x,\theta_i)$,求解下面的最小化问题:

$$\min_{a,b_{st}}\sum_{x=0}^{\infty}w(x)E\Big[\Big(f(x,\theta_i)-a-\sum_{s=1}^{K}\sum_{t=1}^{n}b_{st}I(N_{st}=x)\Big)^2\Big],\tag{1}$$

其中$w(x)$表示已知的权重函数. 得到定理1.

定理1 在前面的假设下,$f(x,\theta_i)$的信度估计为

$$\hat{f}(x,\theta_i)=Zf_i(x)+(1-Z)f_0(x),\tag{2}$$

其中信度因子$Z=n\tau^2/(n\tau^2+\sigma^2)$,而$\tau^2=\sum_{x=0}^{\infty}w(x)\tau^2(x)$,$\sigma^2=\sum_{x=0}^{\infty}w(x)\sigma^2(x)$且$f_i(x)=\dfrac{1}{n}\sum_{j=1}^{n}I(N_{ij}=x)$.

证 由于K个风险之间相互独立,根据信度理论[3],最小化问题(1)等价于

$$\min_{\alpha,\beta_t}\sum_{x=0}^{\infty}w(x)E\Big[\Big(f(x,\theta_i)-\alpha-\sum_{t=1}^{n}\beta_tI(N_{it}=x)\Big)^2\Big].\tag{3}$$

因为

$$E\Big[\Big(f(x,\theta_i)-\alpha-\sum_{t=1}^{n}\beta_tI(N_{it}=x)\Big)^2\Big]=\mathrm{Var}\Big(f(x,\theta_i)-\alpha-\sum_{t=1}^{n}\beta_tI(N_{it}=x)\Big)+\Big[E\Big(f(x,\theta_i)-\alpha-\sum_{t=1}^{n}\beta_tI(N_{it}=x)\Big)\Big]^2\geqslant\mathrm{Var}\Big(f(x,\theta_i)-\alpha-\sum_{t=1}^{n}\beta_tI(N_{it}=x)\Big),$$

上式中等号成立当且仅当$E\Big(f(x,\theta_i)-\alpha-\sum_{t=1}^{n}\beta_tI(N_{it}=x)\Big)=0$,所以$\alpha$的最优选取应为

$$\alpha=E[f(x,\theta_i)]-\sum_{t=1}^{n}\beta_tE[I(N_{it}=x)]=\Big(1-\sum_{t=1}^{n}\beta_t\Big)f_0(x).\tag{4}$$

将(4)式代入(3)式并令

$$\Delta=\sum_{x=0}^{\infty}w(x)E\Big[\Big((f(x,\theta_i)-f_0(x))-\sum_{t=1}^{n}\beta_t(I(N_{it}=x)-f_0(x))\Big)^2\Big].$$

注意到$E[I(N_{it}=x)]=f_0(x)$,并根据双重期望公式有$\mathrm{Var}[I(N_{it}=x)]=\tau^2(x)+\sigma^2(x)$,以及$\mathrm{Cov}[f(x,\theta_i),I(N_{it}=x)]=\tau^2(x)$. 关于$\Delta$对$\beta_k$求导并令导数为零,得到正规方程

$$\sum_{x=0}^{\infty}w(x)\mathrm{Cov}[f(x,\theta_i),I(N_{it}=x)]=\sum_{x=0}^{\infty}w(x)\cdot\mathrm{Cov}\Big(I(N_{ik}=x),\sum_{t=1}^{n}\beta_tI(N_{it}=x)\Big),k=1,2,\cdots,n.$$

解得$\hat{\beta}_k=\tau^2/(n\tau^2+\sigma^2)$,$k=1,2,\cdots,n$,代入(4)式得$\hat{\alpha}=(1-n\tau^2/(n\tau^2+\sigma^2))f_0(x)$. 因此得到(2)式.

注1 不同于文献[8],定理1中的信度因子Z不依赖于x,且信度估计$\hat{f}(x,\theta_i)$是2个概率分布的加权和,且权重满足$0\leqslant Z\leqslant 1$,且$\lim\limits_{n\to\infty}Z=1$. 这恰好遵从信度理论的解释:样本数据越多,则在经验分布函数$f_i(x)$上赋予更多的权重,否则就赋予更少的权重. 这也体现了经验厘定的"奖惩系统"思想.

注2 根据经验分布函数的格列文科定理,容易证明:当$n\to\infty$时有

$$\sup_x|f_i(x)-f(x,\theta_i)|\to 0,\text{a. s.},$$

而$\lim\limits_{n\to\infty}Z=1$,则可得到

$$\sup_x|\hat{f}(x,\theta_i)-f(x,\theta_i)|=\sup_x|Z[f_i(x)-f(x,\theta_i)]+(1-Z)[f_0(x)-f(x,\theta_i)]|\to 0,\text{a. s.},$$

即$\hat{f}(x,\theta_i)$是$f(x,\theta_i)$的强相合估计.

2 3大离散分布中参数的信度估计

在非寿险精算中,刻画离散随机变量的最重要分布包括泊松分布、二项分布和负二项分布. 在这些分布中,将逐个讨论风险参数信度估计.

2.1 泊松分布

假设在θ_i给定时索赔次数N_{ij}服从参数为θ_i的泊松分布,即

$$f(x,\theta_i)=P(N_{ij}=x\mid\theta_i)=\frac{\theta_i^x}{x!}\mathrm{e}^{-\theta_i},x=0,1,2,\cdots,$$

而风险参数θ_i独立同分布于某个先验分布$\pi(\theta)$. 由于此时$\theta_i=E(N_{ij}\mid\theta_i)=\sum_{x=1}^{\infty}xf(x,\theta_i)$. 代入$f(x,\theta_i)$的信度估计可得到$\theta_i$的信度估计为

$$\hat{\theta}_i=\sum_{x=1}^{\infty}x\hat{f}(x,\theta_i)=\sum_{x=1}^{\infty}x(Zf_i(x)+(1-Z)f_0(x))=Z\overline{N}_i+(1-Z)\mu_0,\tag{5}$$

其中 $\overline{N}_i = \frac{1}{n}\sum_{j=1}^{n} N_{ij}$,$\mu_0 = E\theta_i = \int \pi(\theta)\theta \mathrm{d}\theta$,$i = 1,2,\cdots,K$. 若需要估计某个函数 $g(\theta_1,\theta_2,\cdots,\theta_K)$,则有 $\hat{g}(\theta_1,\theta_2,\cdots,\theta_K) = g(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_K)$,其中 $\hat{\theta}_i$ 为(5)式给出.

2.2　二项分布

假设在 θ_i 给定时索赔次数 N_{ij} 服从参数为 θ_i 和 m 的二项分布,即

$$f(x,\theta_i) = P(N_{ij} = x | \theta_i) = \binom{m}{x}\theta_i^x (1-\theta_i)^{m-x},$$

$x = 0,1,2,\cdots,m$,

其中 θ_i 相互独立服从共同的先验分布 $\pi(\theta)$,而 m 已知. 则 $E(N_{ij} | \theta_i) = \sum_{x=1}^{\infty} x f(x,\theta_i) = m\theta_i$,因此 $\theta_i = \frac{1}{m}\sum_{x=1}^{\infty} x f(x,\theta_i)$. 根据 $f(x,\theta_i)$ 的信度估计(2)可得到 θ_i 的信度估计为

$$\hat{\theta}_i = \frac{1}{m}\sum_{x=1}^{\infty} x\hat{f}(x,\theta_i) = Z\frac{1}{m}\overline{N}_i + (1-Z)\frac{\mu_0}{m}, i = 1,2,\cdots,K.$$

类似地,可得到函数 $g(\theta_1,\theta_2,\cdots,\theta_K)$ 的信度估计 $\hat{g}(\theta_1,\theta_2,\cdots,\theta_K) = g(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_K)$.

2.3　负二项分布

假设在 θ_i 给定时索赔次数 N_{ij} 服从参数为 θ_i 和 r 的负二项分布,即

$$f(x,\theta_i) = P(N_{ij} = x | \theta_i) = \binom{x+r-1}{x}\theta_i^x (1-\theta_i)^r, x = 0,1,2,\cdots,$$

其中 θ_i 相互独立服从共同的先验分布 $\pi(\theta)$,而 r 为已知正整数. 注意到负二项分布的条件数学期望为 $E(N_{ij} | \theta_i) = \sum_{x=1}^{\infty} x f(x,\theta_i) = r/(1-\theta_i)$,$i = 1,2,\cdots,K$. 代入 $f(x,\theta_i)$ 的信度估计(2)可解出 θ_i 的信度估计为

$$\hat{\theta}_i = \frac{Z\overline{N}_i + (1-Z)\mu_0 - r}{Z\overline{N}_i + (1-Z)\mu_0}, i = 1,2,\cdots,K,$$

则函数 $g(\theta_1,\theta_2,\cdots,\theta_K)$ 的信度估计 $g(\theta_1,\theta_2,\cdots,\theta_K) = g(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_K)$.

3　结构参数的估计

在实际运用中,$f_0(x)$ 以及结构参数 $\tau^2(x)$、$\sigma^2(x)$ 都是未知的,需要根据样本来进行估计.

首先,给出 $f_0(x)$ 的一个估计 $\hat{f}_0(x) = \frac{1}{K}\sum_{i=1}^{K} f_i(x)$,显然有 $E[\hat{f}_0(x)] = f_0(x)$,即 $\hat{f}_0(x)$ 是 $f_0(x)$ 的无偏估计. 当 $K \to \infty$ 时,根据风险之间的独立性及强大数定律有 $\hat{f}_0(x) \to f_0(x)$,a. s. .

其次,给出结构参数 $\tau^2(x)$ 和 $\sigma^2(x)$ 的无偏估计,即为定理 2.

定理 2　结构参数 $\tau^2(x)$ 和 $\sigma^2(x)$ 的无偏估计分别为

$$\hat{\sigma}^2(x) = \frac{1}{K(n-1)}\sum_{i=1}^{K}\sum_{j=1}^{n}(I(N_{ij} = x) - f_i(x))^2,$$

$$\hat{\tau}^2(x) = \frac{1}{K-1}\sum_{i=1}^{K}(f_i(x) - \hat{f}_0(x))^2 - \frac{1}{n}\hat{\sigma}^2(x).$$

证　记 $S_i = \frac{1}{n-1}\sum_{j=1}^{n}(I(N_{ij} = x) - f_i(x))^2$,$T = \frac{1}{K-1}\sum_{i=1}^{K}(f_i(x) - \hat{f}_i(x))^2$. 由于 $f_i(x) = \frac{1}{n}\sum_{j=1}^{n} I(N_{ij} = x)$,且 $I(N_{ij} = x)$ 在 θ_i 给定下相互独立且具有相同的条件均值和方差

$$E[I(N_{ij} = x) | \theta_i] = f(x,\theta_i),$$

$$\mathrm{Var}[I(N_{ij} = x) | \theta_i] = f(x,\theta_i)(1 - f(x,\theta_i)),$$

则

$$E(S_i) = E[E(S_i | \theta_i)] = E[f(x,\theta_i)(1 - f(x,\theta_i))] = \sigma^2(x).$$

因此 $E[\hat{\sigma}^2(x)] = \frac{1}{K}\sum_{i=1}^{K} E(S_i) = \sigma^2(x)$.

下面证明 $\hat{\tau}^2(x)$ 的无偏性,因为 $E[f_i(x) - \hat{f}_i(x)] = 0$,则有 $E(T) = \frac{1}{K-1}\sum_{i=1}^{K}\mathrm{Var}[f_i(x) - \hat{f}_0(x))]$. 由于

$$\mathrm{Var}[f_i(x) - \hat{f}_i(x))] = \mathrm{Var}(f_i(x)) + \mathrm{Var}(\hat{f}_i(x)) - 2\mathrm{Cov}(f_i(x), \hat{f}_0(x)).$$

根据方差的双重期望公式有

$$\mathrm{Var}(f_i(x)) = \sigma^2(x)/n + \tau^2(x),$$

$$\mathrm{Cov}(f_i(x), \hat{f}_0(x)) = \mathrm{Var}(f_i(x))/K = [\sigma^2(x)/n + \tau^2(x)]/K,$$

$$\mathrm{Var}(\hat{f}_i(x)) = \mathrm{Var}(f_i(x))/K = [\sigma^2(x)/n + \tau^2(x)]/K,$$

则有

$$E(T) = \frac{1}{K-1}\sum_{i=1}^{K}(\sigma^2(x)/n + \tau^2(x))(1 +$$

$1/K - 2/K) = \sigma^2(x)/n + \tau^2(x)$.

因此有 $E[\hat{\tau}^2(x)] = E(T) - E[\hat{\sigma}^2(x)]/n = \tau^2(x)$. 因此 $\hat{\tau}^2(x)$ 是 $\tau^2(x)$ 的无偏估计.

根据 $\tau^2(x)$ 和 $\sigma^2(x)$ 的估计,则得到 τ^2 和 σ^2 的无偏估计.

推论 1 τ^2 和 σ^2 的无偏估计分别为

$$\hat{\sigma}^2 = \frac{1}{K(n-1)}\sum_{x=1}^{\infty} w(x)\sum_{i=1}^{K}\sum_{j=1}^{n}(I(N_{ij} = x) - f_i(x))^2,$$

$$\hat{\tau}^2 = \frac{1}{K-1}\sum_{x=1}^{\infty} w(x)\left[\sum_{i=1}^{K}(f_i(x) - \hat{f}_0(x))^2 - \hat{\sigma}^2(x)/n\right].$$

在实际运用中,有可能出现 $\hat{\tau}^2(x) < 0$ 的情形,这时一般取 $\hat{\tau}^{2*}(x) = \max(0,\hat{\tau}^2(x))$. 但估计 $\hat{\tau}^{2*}(x)$ 不具有无偏性.

4 数值模拟

为了验证本模型与经典的 Bühlmann 信度模型[1] 的区别与联系,给出下面的例子. 由于风险之间相互独立,省略下标 i,即只考虑单个保单的索赔次数分布.

假设在风险参数 θ 给定下,保单在各年索赔次数 N_i 服从泊松分布 Poiss(θ),则 $f(x,\theta) = P(N_i = x \mid \theta) = \theta^x e^{-\theta}/x!, x = 0,1,2,\cdots$, 而 θ 服从 Gamma(α,β) 分布,密度函数为

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}, \theta > 0,$$

则有

$$f_0(x) = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}\frac{\theta^x}{x!}e^{-\theta}d\theta = \frac{\beta^\alpha\Gamma(\alpha+x)}{\Gamma(\alpha)x!(\beta+1)^{x+\alpha}}, x = 0,1,\cdots,$$

$$\sigma^2 = \sum_{x=1}^{\infty} w(x)\left[\frac{\beta^\alpha\Gamma(\alpha+x)}{\Gamma(\alpha)x!(\beta+1)^{x+\alpha}} - \frac{\beta^\alpha\Gamma(\alpha+2x)}{\Gamma(\alpha)(x!)^2(\beta+2)^{2x+\alpha}}\right],$$

$$\tau^2 = \sum_{x=1}^{\infty} w(x)\left[\frac{\beta^\alpha\Gamma(\alpha+2x)}{\Gamma(\alpha)(x!)^2(\beta+2)^{2x+\alpha}} - \frac{\beta^{2\alpha}\Gamma^2(\alpha+x)}{\Gamma^2(\alpha)(x!)^2(\beta+1)^{2x+2\alpha}}\right].$$

因此,$f(x,\theta)$ 的信度估计为

$$\hat{f}(x,\theta) = Zf_n(x) + (1-Z)f_0(x),$$

其中 $Z = n\tau^2/(n\tau^2+\sigma^2)$,而 $f_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i = x)$.

取 $w(x) = I(x = 0)$,$\alpha = 2$,$\beta = 4$,在不同的样本容量 n 以及 θ 值下,计算 $f(x,\theta)$ 及其估计 $\hat{f}(x,\theta)$ 以及相应的均方误差,模拟 10 000 次,得到的结果如表 1 ~ 表 3 所示.

表 1 当 $n = 10$,$\theta = 2$ 时的信度估计及其均方标准差

x	0	1	2	3	4	6	7	⩾8
$f(x,\theta)$	0.135 3	0.270 7	0.270 7	0.180 4	0.090 2	0.036 1	0.012 0	0.004 5
$\hat{f}(x,\theta)$	0.166 7	0.276 9	0.255 2	0.170 4	0.078 4	0.030 5	0.012 3	0.009 6
均方标准差	0.103 7	0.130 1	0.129 1	0.109 6	0.081 0	0.054 2	0.033 1	0.021 9

表 2 当 $n = 50$,$\theta = 2$ 时的信度估计及其均方标准差

x	0	1	2	3	4	6	7	⩾8
$f(x,\theta)$	0.135 3	0.270 7	0.270 7	0.180 4	0.090 2	0.036 1	0.012 0	0.004 5
$\hat{f}(x,\theta)$	0.143 9	0.269 7	0.269 6	0.176 2	0.087 8	0.036 1	0.011 5	0.005 2
均方标准差	0.046 5	0.063 6	0.061 3	0.053 7	0.038 3	0.025 6	0.014 9	0.008 8

表 3 当 $n = 200$,$\theta = 2$ 时的信度估计及其均方标准差

x	0	1	2	3	4	6	7	⩾8
$f(x,\theta)$	0.135 3	0.270 7	0.270 7	0.180 4	0.090 2	0.036 1	0.012 0	0.004 5
$\hat{f}(x,\theta)$	0.139 2	0.271 8	0.268 3	0.180 0	0.088 1	0.036 3	0.011 5	0.004 8
均方标准差	0.024 6	0.032 2	0.031 8	0.027 1	0.020 4	0.013 2	0.007 2	0.004 9

在表 1 ~ 表 3 中,$f(x,\theta)$ 为待估的点概率,而 $\hat{f}(x,\theta)$ 为 10 000 次模拟中信度估计的均值. 从模拟结果可以看出,信度估计 $\hat{f}(x,\theta)$ 收敛于点概率 $f(x,\theta)$,即使样本量较小($n = 50$),其均方标准差已经控制在 5% 以下. 为了体现模型的优势,分别计算风险参数 θ 的信度估计. 根据信度估计 $\hat{f}(x,\theta)$ 可得到风险参数 θ 的信度估计为 $\hat{\theta} = Z\overline{N} + (1-Z)\mu_0$,其中

$\overline{N}=\sum_{i=1}^{n}N_i$,信度因子 $Z=n\tau^2/(n\tau^2+\sigma^2)$ 为(2)式给出. 而文献[1]的信度估计为 $\hat{\theta}^B=Z^B\overline{N}+(1-Z^B)\mu_0$,这里

$$Z^B=n\mathrm{Var}(E(N|\theta))/[n\mathrm{Var}(E(N|\theta))+E(\mathrm{Var}(N|\theta))]=(n\alpha/\beta^2)/(n\alpha/\beta^2+\alpha/\beta)=n/(n+\beta).$$

将本文的信度估计与经典的 Bühlmann 信度估计的均方标准差进行比较,得到表 4.

表 4　当 $\theta=0.8$ 时风险参数的信度估计的均方标准差的比较

n	10	30	80	150	300	800	2 000
$S(\hat{\theta})$	0.262 4	0.151 8	0.095 2	0.071 4	0.052 3	0.036 8	0.028 8
$S(\hat{\theta}^B)$	0.219 4	0.143 5	0.110 6	0.100 3	0.092 9	0.088 7	0.086 8

在表4中,$S(\hat{\theta})$ 及 $S(\hat{\theta}^B)$ 分别表示本文给出的信度估计及 Bühlmann 信度估计的均方标准差. 从表 4 可以看出,当样本容量较小时,本文给出的信度估计 θ 的均方差比 Bühlmann 信度估计稍大一些,然而,随着样本容量的增大,信度估计 $\hat{\theta}$ 的均方差比 Bühlmann 信度估计的均方误差更小. 分析其原因,2 个信度估计的均方误差大小的差异主要体现在信度因子不同导致的,本文给出的信度因子以更快的速度收敛于 1,因此当样本容量较大时有较小的均方误差. 总体来说,2 个信度估计均有较小的均方误差,能满足实际使用的需要.

5　参考文献

[1] Bühlmann H. Experience rating and credibility[J]. Astin Bulletin,1967,4(3):199-207.

[2] Bühlmann H,Gisler A. A course in credibility theory and its applications[M]. Netherlands:Springer,2005.

[3] Norberg R. Credibility theory[M]. Chichester:John Wiley and Sons,2006:81-156.

[4] Mowbray A H. How extensive a payroll is necessary to give a dependable pure premium?[J]. Proceedings of the Casualty Actuarial Society,1914(1):24-30.

[5] Whitney A. The theory of experience rating[J]. Proceedings of the Casualty Actuarial Society,1918(4):274-292.

[6] 杨静平. 非寿险精算学[M]. 北京:北京大学出版社,2006.

[7] 温利民. 信度估计的理论与方法[M]. 北京:科学出版社,2012.

[8] Pan Maolin,Wang Rongming,Wu Xianyi. On the consistency of credibility premiums regarding esscher principle[J]. Insurance: Mathematics and Economics,2008,42(1):119-126.

[9] Gerber H U. Credibility for esscher premium[J]. Mitteilungen der Vereinigung schweiz Versicher ungsmathematiker,Heft,1980(3):307-312.

[10] Wen Linming,Wu Xianyi. Experience rating under the exponential premium principle[J]. Sci Sin Math,2011,41(10):861-876.

[11] 张林娜,温利民,方婧. 聚合风险模型下指数保费的非参数估计[J]. 西南大学学报:自然科学版,2016,38(5):100-105.

[12] 温利民,庄小红. 零期望效用原理下的贝叶斯保费[J]. 系统科学与数学,2016,36(8):1318-1328.

The Weighted Credibility Estimation and Their Statistical Analysis of Claims Number

ZHANG Yi,LI Zhilong*

(School of Statistics,Jiangxi University of Finance and Economics,Nanchang Jiangxi 330013,China)

Abstract: Using the idea of credibility theory,the credibility estimation of claim number is studied,the credibility estimation of the probability distribution of the claim number is obtained by minimizing the expected weighted sum loss function. Furthermore,credibility estimates of risk parameters are discussed respectively in the three discrete distributions of claim number,and the estimation of structural parameters are obtained. The unbiasedness of estimators are proved. Finally,the credibility estimator is compared with that of the classical Bühlmann estimator by numerical simulation,which verifies the convergence rate of the proposed estimator.

Key words: claims number;risk parameter;credibility estimate;structural parameter

(责任编辑: 曾剑锋)