应用概率统计 第二十七卷
第三期 2011年6月

Chinese Journal of Applied Probability
and Statistics Vol.27 No.3 Jun. 2011

具有共同效应的信度回归模型 *

王筑娟
(上海应用技术学院数理部, 上海, 200235)

温利民
(江西师范大学数信学院, 南昌, 330022)

摘 要

在经典的Hachemeister (1975)信度回归模型中, 各个风险被假定为相互独立的. 本文假设风险之间存在由共同效应导致的风险相依, 建立了共同效应的信度回归模型, 得到未来索赔的信度预测与风险参数的信度估计. 结论表明, 在共同效应模型, 信度估计仍然是个体索赔数据与聚合保费的加权和.

关键词: 信度估计, 共同效应, 正交投影, 回归信度.

学科分类号: O211.9.

§1. 引 言

信度回归模型是Hachemeister (1975)[1]提出的模型. Hachemeister在利用Bülmann-Straub信度模型对美国各州的汽车第三者责任险进行定价时, 注意到索赔数据在时间分量上由于通货膨胀的影响具有时间趋势效应, 提出提出用回归模型来刻画该效应, 建立了信度回归模型. 信度回归模型可以简单刻画如下: 假设在第i份风险合同下, 当风险参数β_i给定时, 对索赔X_i与未来索赔X_{i,n_i+1}有下面的模型成立:

$$X_i = Y_i\beta_i + \epsilon_i \tag{1.1}$$

以及

$$X_{i,n_i+1} = Y_{i,n_i+1}\beta_i + \epsilon_{i,n_i+1}. \tag{1.2}$$

这里$i=1,2,\ldots,K$, $X_i=(X_{i1},\ldots,X_{in_i})'$为索赔样本, 其中变量$Y_i, Y_{i,n_i+1}$为已知的设计向量, 而$\beta_i$ $(i=1,2,\ldots,K)$为未知不可观测的随机系数向量, ϵ_i $(i=1,2,\ldots,K)$为随机误差向量, 也是不可观测的. 在对未来索赔$X_{i,n_{i+1}}$进行预测时, 假设K个风险之间是相互独立的. 然而, 风险之间的独立性假设仅仅是实际风险的一种理想化近似. 众所周知, 大多情况下风险之间都存在相依性. 例如, 研究表明, 夫妻之间的寿命呈现正相依性, 在相同地域下的承保房屋之间面临共同的火灾风险; 一次交通事故可能导致多个承保车辆受损, 等等. 近年来, 对风险之间的相依性的研究受到了越来越多的精算研究者的关注. 例如[2]-[5]. 注意到, Yeo等(2006)[6]提出一种共同效应(common effect)随机模型, 用来刻画风险之间的相依

*上海应用技术学院科学技术发展基金项目(KJ2008-15)、江西省教育厅青年科学基金项目(GJJ10096)资助.
本文2010年7月22日收到.

性, 他们获得了正态–正态分布下的信度公式, Wen等(2009)[7]将该模型推广无分布(任意分布)的情形, 得到风险在共同效应相依情况下的信度保费. 本文将考虑风险之间存在共同效应时的信度回归模型.

本文分为四节, 第二节对矩阵投影以及信度理论作一介绍. 第三节建立共同效应信度回归模型, 得到该模型下未来索赔的信度预测与风险参数的信度估计, 最后一节是全文的总结.

§2. 模型假设与准备知识

考虑K份保险合同, 在精算学中也表示K个风险, 我们将研究这K个风险之间由于某种共同效应导致相依时的回归信度估计. 在信度模型中, 假设第i个风险由风险参数β_i识别, 这里β_i是对该风险的风险属性, 例如汽车保险中的驾驶人的性别、年龄、个人爱好、工作性质等的综合刻画. 在信度理论中, 假设风险参数为不可观测的随机向量. 对每个保险合同, 已经有若干年的索赔历史记录, 这些记录在统计中称为样本. 在信度回归模型中, 假设索赔随机变量服从线性回归模型. 基于这些索赔样本, 我们的目标是预测每个合同在未来年的索赔或估计风险参数. 由于风险之间存在共同效应, 因此, 这K个风险保险合同之间具有某种相依性. 把这个模型的基本假设列于下面的假设2.1-2.3.

假设 2.1 共同效应变量Λ具有已知的期望$\mathsf{E}(\Lambda)=\mu_\lambda$和方差$\mathsf{Var}\,(\Lambda)=\sigma_\lambda^2$.

假设 2.2 给定共同效应$\Lambda=\lambda$时, 假设风险参数Θ_i, $i=1,\ldots,K$为独立同分布的随机变量序列, 其共同的分布函数为$\pi(\theta|\lambda)$.

假设 2.3 对固定的风险i, 索赔样本$X_i=(X_{i1},X_{i2},\ldots,X_{i,n_i+1})'$与未来索赔$X_{i,n_i+1}$服从线性回归模型

$$X_i=Y_i\beta(\Theta_i,\Lambda)+\varepsilon_i \tag{2.1}$$

以及

$$X_{i,n_i+1}=Y_{i,n_i+1}\beta(\Theta_i,\Lambda)+\varepsilon_{i,n_i+1}. \tag{2.2}$$

这里

- $Y_i=(Y_{ij})_{n_i\times p}$为$n_i\times p$维的设计矩阵, 表示风险$i$在前$n_i$年的解释变量的观察值, Y_{i,n_i+1}为设计变量, 表示在未来一年的观察变量的取值, 均为已知;

- $(\varepsilon_i',\varepsilon_{i,n_i+1})$为具有零均值误差项. 假设当$\Lambda$给定时, $(\varepsilon_i',\varepsilon_{i,n_i+1})$, $i=1,\ldots,K$是相互独立的, 且其条件方差或协方差为

$$\mathsf{Cov}\,(\varepsilon_i,\varepsilon_i|\Theta_i,\Lambda)=\sigma^2(\Theta_i,\Lambda)V_i \quad 以及 \quad \mathsf{Cov}\,(\varepsilon_{i,n_i+1},\varepsilon_i)=\sigma^2(\Theta_i,\Lambda)R_i, \tag{2.3}$$

这里, $\sigma^2(\Theta_i,\Lambda)>0$是$\Theta_i$和$\Lambda$的未知函数, V_i和R_i是已知的矩阵或向量. 最简单的情形是$V_i=I_{n_i}$ (n_i阶的单位矩阵)和$R_i=(0,\ldots,0)$ (n_i阶的行向量).

为叙述的方便, 记

$$\begin{aligned}\beta_1(\Lambda) &= \mathsf{E}(\beta(\Theta_i,\Lambda)|\Lambda),\\ \sigma_1^2(\Lambda) &= \mathsf{E}(\sigma^2(\Theta_i,\Lambda)|\Lambda),\\ N(\Lambda) &= \mathsf{Var}\,(\beta(\Theta_i,\Lambda)|\Lambda),\end{aligned}\tag{2.4}$$

以及

$$\begin{aligned}&\beta_0 = \mathsf{E}[\beta_1(\Lambda)],\\ &\sigma_0^2 = \mathsf{E}[\sigma_1^2(\Lambda)], \qquad \mathsf{Var}\,[\beta_1(\Lambda)] = M,\\ &N = \mathsf{E}[N(\Lambda)].\end{aligned}\tag{2.5}$$

显然, $\mathsf{Var}\,(\beta(\Theta_i,\Lambda)) = M+N$, $\mathsf{Cov}\,(\beta(\Theta_i,\Lambda),\beta(\Theta_j,\Lambda)) = M$, $i\neq j$.

在该共同效应信度回归模型中, 我们的目标是基于样本数据(X_i,Y_i,Y_{i,n_i+1}), $i=1,2,\ldots,K$预测未来的索赔X_{i,n_i+1}. 信度保费是将预测函数限定在样本的线性函数类中, 使得期望平方损失函数达到最小, 即我们将求解下面的最优化问题:

$$\min_{a_0,a_{st}\in R}\mathsf{E}\Big[\Big(X_{i,n_i+1}-a_0-\sum_{s=1}^{K}\sum_{t=1}^{n_s}a_{st}X_{st}\Big)^2\Big].\tag{2.6}$$

下面的引理可以更方便地求信度估计, 其证明可参考Wen等(2009)[7].

引理 2.1 设随机向量$\begin{pmatrix}X_{p\times1}\\ Y_{q\times1}\end{pmatrix}$的期望与方差协方差矩阵分别为

$$\begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix} \quad 与 \quad \begin{pmatrix}\Sigma_{XX} & \Sigma_{XY}\\ \Sigma_{YX} & \Sigma_{YY}\end{pmatrix}.$$

则当

$$A=\mu_Y-\Sigma_{YX}\Sigma_{XX}^{-1}\mu_X B=\Sigma_{YX}\Sigma_{XX}^{-1}\tag{2.7}$$

时, 期望损失

$$\mathsf{E}(Y-A-BX)(Y-A-BX)'\tag{2.8}$$

在矩阵的非负定意义下达到最小.

因此, 根据引理2.1, 基于随机向量X的非齐次函数类的随机向量Y的最优预测为

$$\mathrm{proj}(Y|L(X,1))=\mu_Y+\Sigma_{YX}\Sigma_{XX}^{-1}(X-\mu_X),\tag{2.9}$$

其中$\mathrm{proj}(Y|L(X,1))$表示Y在X的线性函数空间的正交投影. 表达式(2.9)称为随机向量Y在X的线性空间的(非齐次)正交投影.

对正交投影算子“proj”, 有下面的结论.

引理 2.2 对L^2的两个闭子空间M, M', 若$M' \subset M \subset L^2$, 设随机向量$Y \in L^2$, 则有

$$\mathrm{proj}(X|M') = \mathrm{proj}(\mathrm{proj}(X|M)|M'). \tag{2.10}$$

式(2.10)被称为正交投影的平滑性, 其证明可参考[8].

下面给出一个矩阵求逆公式, 其证明可参考[9].

引理 2.3 设A, B, C, D是适当阶数的矩阵, 则有下面的求逆公式

$$(A+BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}. \tag{2.11}$$

§3. 具有共同效应的回归信度估计

根据引理2.1和引理2.2, 容易得到X_{i,n_i+1}的信度保费$\widehat{X}^*_{i,n_i+1}$.

定理 3.1 在假设2.1-2.3成立时, 第i个合同的信度保费$\widehat{X}^*_{i,n_i+1}$为

$$\widehat{X}^*_{i,n_i+1} = R_iV_i^{-1}X_i + (\lambda_i - R_iV_i^{-1})Y_i\widehat{\beta(\Theta_i,\Lambda)}^*, \tag{3.1}$$

其中λ_i为线性方程

$$Y_{i,n_i+1} = \lambda_i Y_i \tag{3.2}$$

的解, 且$Y_i\widehat{\beta(\Theta_i,\Lambda)}^* = \mathrm{proj}(Y_i\beta(\Theta_i,\Lambda)|L(X,1))$为$Y_i\beta(\Theta_i,\Lambda)$的信度估计.

证明: 根据引理2.1, 第i个合同的信度保费为

$$\widehat{X}^*_{i,n_i+1} = \mathrm{proj}(X_{i,n_i+1}|L(X,1)). \tag{3.3}$$

由于$L(X,1) \subseteq L(X,\Theta,\Lambda,1)$, 则正交投影的平滑性(2.10), 有

$$\widehat{X}^*_{i,n_i+1} = \mathrm{proj}(\mathrm{proj}(X_{i,n_i+1}|L(X,\Theta,\Lambda,1))|L(X,1)),$$

其中$X = (X_1', X_2', \ldots, X_K')'$, 这里"$'$"表示矩阵的转置. 又因为

$$\begin{aligned} & \mathrm{proj}(X_{i,n_i+1}|L(X,\Theta,\Lambda,1)) \\ = \ & \mathsf{E}(X_{i,n_i+1}|\Theta,\Lambda) + \mathsf{Cov}\,(X_{i,n_i+1}, X|\Theta,\Lambda)\mathsf{Cov}\,(X,X|\Theta,\Lambda)^{-1}(X - \mathsf{E}(X|\Theta,\Lambda)). \end{aligned}$$

根据假设2.1-2.3, 有

$$\mathsf{E}(X_{i,n_i+1}|\Theta,\Lambda) = Y_{i,n+1}\beta(\Theta_i,\Lambda), \qquad \mathsf{Cov}\,(X_{i,n_i+1}, X|\Theta,\Lambda) = e_i \otimes R_i\sigma^2(\Theta_i,\Lambda)$$

以及

$$\mathrm{Cov}\,(X,X|\Theta,\Lambda)^{-1} = \begin{pmatrix} V_1\sigma^2(\Theta_1,\Lambda) & & \\ & \ddots & \\ & & V_K\sigma^2(\Theta_K,\Lambda) \end{pmatrix}^{-1} = \begin{pmatrix} \dfrac{V_1^{-1}}{\sigma^2(\Theta_1,\Lambda)} & & \\ & \ddots & \\ & & \dfrac{V_K^{-1}}{\sigma^2(\Theta_K,\Lambda)} \end{pmatrix}.$$

因此

$$\begin{aligned}\mathrm{proj}(X_{i,n_i+1}|L(X,\Theta,\Lambda,1)) &= Y_{i,n_i+1}\beta(\Theta_i,\Lambda)+R_iV_i^{-1}(X_i-Y_i\beta(\Theta_i,\Lambda))\\ &= Y_{i,n_i+1}\beta(\Theta_i,\Lambda)+R_iV_i^{-1}(X_i-Y_i\beta(\Theta_i,\Lambda))\\ &= R_iV_i^{-1}X_i+(Y_{i,n_i+1}-R_iV_i^{-1}Y_i)\beta(\Theta_i,\Lambda)\\ &= R_iV_i^{-1}X_i+(\lambda_i-R_iV_i^{-1})Y_i\beta(\Theta_i,\Lambda).\end{aligned}$$

又因为$R_iV_i^{-1}X_i\in L(X,1)$, 则

$$\begin{aligned}\widehat{X}^*_{i,n+1} &= \mathrm{proj}(R_iV_i^{-1}X_i+(\lambda_i-R_iV_i^{-1})Y_i\beta(\Theta_i,\Lambda)|L(X,1))\\ &= R_iV_i^{-1}X_i+(\lambda_i-R_iV_i^{-1})\mathrm{proj}(Y_i\beta(\Theta_i,\Lambda)|L(X,1)) &(3.4)\\ &= R_iV_i^{-1}X_i+(\lambda_i-R_iV_i^{-1})Y_i\widehat{\beta(\Theta_i,\Lambda)}^*. &(3.5)\end{aligned}$$

定理证完. $\square$

根据定理3.1, 若要求得信度保费估计$\widehat{X}^*_{i,n_i+1}$, 则只需求$Y_i\beta(\Theta_i,\Lambda)$的信度估计$Y_i\widehat{\beta(\Theta_i,\Lambda)}^*$. 首先给出各变量之间的协方差关系.

引理 3.1 在假设2.1-2.3成立时, 有

(1) 随机向量$Y_i\beta(\Theta_i,\Lambda)$与随机向量$X$之间的协方差矩阵为

$$\mathrm{Cov}\,(Y_i\beta(\Theta_i,\Lambda),X)=Y_iM(Y_1',Y_2',\ldots,\ldots,Y_K')+Y_iN(0,\ldots,Y_i',\ldots,0). \tag{3.6}$$

(2) 随机向量X的方差协方差矩阵为

$$\mathrm{Cov}\,(X,X)=\begin{pmatrix}Y_1\\ \vdots\\ Y_K\end{pmatrix}M\begin{pmatrix}Y_1' & \cdots & Y_K'\end{pmatrix}+\begin{pmatrix}U_1 & & \\ & \ddots & \\ & & U_k\end{pmatrix}, \tag{3.7}$$

其中$U_i=\sigma_0^2V_i+Y_iNY_i'$.

(3) 协方差矩阵$\mathrm{Cov}\,(X,X)$的逆矩阵为

$$\begin{aligned}\mathrm{Cov}\,(X,X)^{-1} &= \begin{pmatrix} U_1^{-1} & & \\ & \ddots & \\ & & U_K^{-1}\end{pmatrix}\\ &\quad - \begin{pmatrix} U_1^{-1}Y_1 \\ \vdots \\ U_K^{-1}Y_K\end{pmatrix}(M^{-1}+\Delta)^{-1}\begin{pmatrix} Y_1'U_1^{-1} & \cdots & Y_K'U_1^{-1}\end{pmatrix}, \end{aligned} \tag{3.8}$$

其中

$$\Delta = \sum_{s=1}^{K} Y_s'(\sigma_0^2 V_s + Y_s N Y_s')^{-1} Y_s. \tag{3.9}$$

证明: 首先, 注意到

$$\begin{aligned}\mathrm{Cov}\,(Y_i\beta(\Theta_i,\Lambda),X_j) &= \mathrm{E}[\mathrm{Cov}\,(Y_i\beta(\Theta_i,\Lambda),X_j|\Theta_i,\Lambda)] + \mathrm{Cov}\,(Y_i\beta(\Theta_i,\Lambda),Y_j\beta(\Theta_j,\Lambda))\\ &= \begin{cases} Y_iMY_i'+Y_iNY_i', & i=j;\\ Y_iMY_j', & i\neq j.\end{cases}\end{aligned}$$

则有

$$\begin{aligned}\mathrm{Cov}\,(Y_i\beta(\Theta_i,\Lambda),X) &= (Y_iMY_1',\ldots,Y_iMY_i'+Y_iNY_i',\ldots,Y_iMY_K')\\ &= (Y_iMY_1',Y_iMY_2',\ldots,Y_iMY_i',\ldots,Y_iMY_K')\\ &\quad +(0,\ldots,Y_iNY_i',\ldots,0)\\ &= Y_iM(Y_1',Y_2',\ldots,\ldots,Y_K')+Y_iN(0,\ldots,Y_i',0).\end{aligned}$$

因此

$$\begin{aligned}\mathrm{Cov}\,(X_i,X_j) &= \mathrm{E}[\mathrm{Cov}\,(X_i,X_j|\Theta_i,\Lambda)] + \mathrm{Cov}\,(Y_i\beta(\Theta_i,\Lambda),Y_j\beta(\Theta_j,\Lambda))\\ &= \begin{cases} Y_i(M+N)Y_i'+\sigma_0^2V_i, & i=j;\\ Y_iMY_j', & i\neq j.\end{cases}\end{aligned}$$

所以得到

$$\begin{aligned}\Sigma_{XX} &= \begin{pmatrix} Y_1(M+N)Y_1'+\sigma_0^2V_1 & \cdots & Y_1MY_K'\\ \cdots & \cdots & \cdots\\ Y_KMY_1' & \cdots & Y_K(M+N)Y_K'+\sigma_0^2V_K\end{pmatrix}\\ &= \begin{pmatrix} Y_1\\ \vdots\\ Y_K\end{pmatrix} M \begin{pmatrix} Y_1' & \cdots & Y_K'\end{pmatrix} + \begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_K\end{pmatrix}.\end{aligned}$$

由矩阵求逆公式(2.11)有

$$\begin{aligned}\mathsf{Cov}\,(X,X)^{-1} &= \begin{pmatrix} U_1^{-1} & & \\ & \ddots & \\ & & U_K^{-1}\end{pmatrix} \\ &\quad - \begin{pmatrix} U_1^{-1}Y_1 \\ \vdots \\ U_K^{-1}Y_K\end{pmatrix}(M^{-1}+\Delta)^{-1}\begin{pmatrix} Y_1'U_1^{-1} & \cdots & Y_K'U_1^{-1}\end{pmatrix}.\end{aligned}$$

则完成了引理的证明. □

记

$$\widehat{\beta}_{\mathrm{GL}(s)} = (Y_s'V_s^{-1}Y_s)^{-1}Y_s'V_s^{-1}X_s$$

表示假设2.3中第s个线性回归方程的最小二乘解, 而

$$\widehat{\beta}_{\mathrm{GL}} = \sum_{r=1}^{K}\Delta^{-1}\Delta_r\widehat{\beta}_{\mathrm{GL}(r)}$$

为这K个最小二乘解的加权和, 其中权重为

$$\Delta_s = Y_s'(\sigma_0^2V_s + Y_sNY_s')^{-1}Y_s, \qquad s=1,2,\ldots,K.$$

定理 3.2 在假设2.1-2.3成立时, $Y_i\beta(\Theta_i,\Lambda)$的信度估计$Y_i\widehat{\beta(\Theta_i,\Lambda)}^*$为

$$Y_i\widehat{\beta(\Theta_i,\Lambda)}^* = Y_i(Z_{i1}\widehat{\beta}_{\mathrm{GL}(i)} + Z_{i2}\widehat{\beta}_{\mathrm{GL}} + (I_p - Z_{i1} - Z_{i2})\beta_0), \tag{3.10}$$

其中信度因子Z_{i2}与Z_{i2}分别为

$$Z_{i1} = N\Delta_i, \qquad Z_{i2} = M(M+\Delta^{-1})^{-1} - N\Delta_i(M^{-1}+\Delta)^{-1}\Delta. \tag{3.11}$$

证明: 根据(2.9)式, 有

$$\begin{aligned}Y_i\widehat{\beta(\Theta_i,\Lambda)}^* &= \mathsf{E}(\beta(\Theta_i,\Lambda)) + \mathsf{Cov}\,(Y_i\beta(\Theta_i,\Lambda),X)\mathsf{Cov}\,(X,X)^{-1}(X-\mathsf{E}X) \\ &= Y_i\beta_0 + \mathsf{Cov}\,(Y_i\beta(\Theta_i,\Lambda),X)\mathsf{Cov}\,(X,X)^{-1}\begin{pmatrix} X_1 - Y_1\beta_0 \\ \vdots \\ X_K - Y_K\beta_0\end{pmatrix}.\end{aligned} \tag{3.12}$$

注意到

$$Y_s'V_s^{-1}X_s = Y_s'V_s^{-1}Y_s[(Y_s'V_s^{-1}Y_s)^{-1}Y_s'V_s^{-1}X_s] = Y_s'V_s^{-1}Y_s\widehat{\beta}_{\mathrm{GL}(s)}. \tag{3.13}$$

应用矩阵求逆公式(2.11), 则有

$$
\begin{aligned}
& Y_s'(\sigma_0^2V_s+Y_sNY_s')^{-1}X_s \\
= \ & Y_s'\Big(\frac{1}{\sigma_0^2}V_s^{-1}-\frac{1}{\sigma_0^4}V_s^{-1}Y_s\Big(N^{-1}+\frac{1}{\sigma_0^2}Y_s'V_s^{-1}Y_s\Big)^{-1}Y_s'V_s^{-1}\Big)X_s \\
= \ & \Big(\frac{1}{\sigma_0^2}I_n-\frac{1}{\sigma_0^4}Y_s'V_s^{-1}Y_s\Big(N^{-1}+\frac{1}{\sigma_0^2}Y_s'V_s^{-1}Y_s\Big)^{-1}\Big)Y_s'V_s^{-1}X_s \\
= \ & \Big(\frac{1}{\sigma_0^2}I_n-\frac{1}{\sigma_0^4}Y_s'V_s^{-1}Y_s\Big(N^{-1}+\frac{1}{\sigma_0^2}Y_s'V_s^{-1}Y_s\Big)^{-1}\Big)Y_s'V_s^{-1}Y_s\widehat{\beta}_{\mathrm{GL}(s)} \\
= \ & Y_s'\Big(\frac{1}{\sigma_0^2}V_s^{-1}-\frac{1}{\sigma_0^4}V_s^{-1}Y_s\Big(N^{-1}+\frac{1}{\sigma_0^2}Y_s'V_s^{-1}Y_s\Big)^{-1}Y_s'V_s^{-1}\Big)Y_s\widehat{\beta}_{\mathrm{GL}(s)} \\
= \ & Y_s'(\sigma_0^2V_s+Y_sNY_s')^{-1}Y_s\widehat{\beta}_{\mathrm{GL}(s)}.
\end{aligned}
\tag{3.14}
$$

注意到(3.6)与(3.7), 结合式(3.13)和(3.14)有

$$
\begin{aligned}
& Y_iM(Y_1',Y_2',\ldots,\ldots,Y_K')\begin{pmatrix} U_1^{-1} & & \\ & \ddots & \\ & & U_K^{-1}\end{pmatrix}\begin{pmatrix} X_1-Y_1\beta_0 \\ \vdots \\ X_K-Y_K\beta_0\end{pmatrix} \\
= \ & \sum_{s=1}^K Y_iMY_s'(\sigma_0^2V_s+Y_sNY_s')^{-1}(X_s-Y_s\beta_0) \\
= \ & Y_iM\Big[\sum_{s=1}^K Y_s(\sigma_0^2V_s+Y_sNY_s')^{-1}X_s-\sum_{s=1}^K Y_s(\sigma_0^2V_s+Y_sNY_s')^{-1}Y_s\beta_0\Big] \\
= \ & Y_iM\Big[\sum_{s=1}^K \Delta_s(\widehat{\beta}_{\mathrm{GL}(s)}-\beta_0)\Big],
\end{aligned}
\tag{3.15}
$$

并且

$$
\begin{aligned}
& Y_iM(Y_1',Y_2',\ldots,\ldots,Y_K')\begin{pmatrix} U_1^{-1}Y_1 \\ \vdots \\ U_K^{-1}Y_K\end{pmatrix}(M^{-1}+\Delta)^{-1} \\
& \cdot\begin{pmatrix} Y_1'U_1^{-1} & \cdots & Y_K'U_1^{-1}\end{pmatrix}\begin{pmatrix} X_1-Y_1\beta_0 \\ \vdots \\ X_K-Y_K\beta_0\end{pmatrix} \\
= \ & Y_iM\Big(\sum_{s=1}^K\Delta_s\Big)(M^{-1}+\Delta)^{-1}\Big(\sum_{r=1}^K\Delta_r(\widehat{\beta}_{\mathrm{GL}(r)}-\beta_0)\Big) \\
= \ & Y_iM\Delta(M^{-1}+\Delta)^{-1}\Big(\sum_{r=1}^K\Delta_r(\widehat{\beta}_{\mathrm{GL}(r)}-\beta_0)\Big),
\end{aligned}
\tag{3.16}
$$

以及

$$Y_iN(0,\ldots,Y_i',\ldots,0)\begin{pmatrix}U_1^{-1}&&\\&\ddots&\\&&U_K^{-1}\end{pmatrix}\begin{pmatrix}X_1-Y_1\beta_0\\ \vdots\\ X_K-Y_K\beta_0\end{pmatrix}=Y_iN\Delta_i(\widehat{\beta}_{\mathrm{GL}(i)}-\beta_0). \tag{3.17}$$

又因为

$$\begin{aligned}
&Y_iN(0,\ldots,Y_i',\ldots,0)\begin{pmatrix}U_1^{-1}Y_1\\ \vdots\\ U_K^{-1}Y_K\end{pmatrix}(M^{-1}+\Delta)^{-1}\\
&\cdot\begin{pmatrix}Y_1'U_1^{-1}&\cdots&Y_K'U_1^{-1}\end{pmatrix}\begin{pmatrix}X_1-Y_1\beta_0\\ \vdots\\ X_K-Y_K\beta_0\end{pmatrix}\\
=\;&Y_iNY_i'(\sigma_0^2V_i+Y_iNY_i')^{-1}Y_i(M^{-1}+\Delta)^{-1}\Big(\sum_{r=1}^{K}Y_rMY_r'(\sigma_0^2V_r+Y_rNY_r')^{-1}(X_r-Y_r\beta_0)\Big)\\
=\;&Y_iN\Delta_i(M^{-1}+\Delta)^{-1}\Big(\sum_{r=1}^{K}\Delta_r(\widehat{\beta}_{\mathrm{GL}(r)}-\beta_0)\Big).
\end{aligned} \tag{3.18}$$

则将(3.15)-(3.18)代入(3.12), 得到$Y_i\beta(\Theta_i,\Lambda)$的信度估计为

$$\begin{aligned}
Y_i\widehat{\beta(\Theta_i,\Lambda)}^* &= Y_i\beta_0+Y_iM\Big[\sum_{s=1}^{K}\Delta_s(\widehat{\beta}_{\mathrm{GL}(s)}-\beta_0)\Big]\\
&\quad -Y_iM\Delta(M^{-1}+\Delta)^{-1}\Big(\sum_{r=1}^{K}\Delta_r(\widehat{\beta}_{\mathrm{GL}(r)}-\beta_0)\Big)\\
&\quad +Y_iN\Delta_i(\widehat{\beta}_{\mathrm{GL}(i)}-\beta_0)-Y_iN\Delta_i(M^{-1}+\Delta)^{-1}\Big(\sum_{r=1}^{K}\Delta_r(\widehat{\beta}_{\mathrm{GL}(r)}-\beta_0)\Big)\\
&= Y_i\beta_0+Y_iZ_{i1}(\widehat{\beta}_{\mathrm{GL}(i)}-\beta_0)+Y_iZ_{i2}\sum_{r=1}^{K}\Delta^{-1}\Delta_r(\widehat{\beta}_{\mathrm{GL}(r)}-\beta_0)\\
&= Y_i(Z_{i1}\widehat{\beta}_{\mathrm{GL}(i)}+Z_{i2}\widehat{\beta}_{\mathrm{GL}}+(I_p-Z_{i1}-Z_{i2})\beta_0). \qquad\Box
\end{aligned}$$

注记 1 根据定理3.2, 由于

$$Y_i\widehat{\beta(\Theta_i,\Lambda)}^*=Y_i(Z_{i1}\widehat{\beta}_{\mathrm{GL}(i)}+Z_{i2}\widehat{\beta}_{\mathrm{GL}}+(I_p-Z_{i1}-Z_{i2})\beta_0).$$

因此, 记

$$\widehat{\beta(\Theta_i,\Lambda)}^*=Z_{i1}\widehat{\beta}_{\mathrm{GL}(i)}+Z_{i2}\widehat{\beta}_{\mathrm{GL}}+(I_p-Z_{i1}-Z_{i2})\beta_0. \tag{3.19}$$

这说明, 风险参数$\beta(\Theta_i,\Lambda)$的信度估计$\widehat{\beta(\Theta_i,\Lambda)}^*$可以表达为第$i$各合同的最小二乘估计$\widehat{\beta}_{\mathrm{GL}(i)}$, 总最小二乘估计$\widehat{\beta}_{\mathrm{GL}}$以及聚合均值$\beta_0$的加权平均. 事实上, 类似地可以证明对任

意$1\times p$的常数向量C, 有

$$C\widehat{\beta(\Theta_i,\Lambda)}^* = C[Z_{i1}\widehat{\beta}_{\mathrm{GL}(i)} + Z_{i2}\widehat{\beta}_{\mathrm{GL}} + (I_p - Z_{i1} - Z_{i2})\beta_0].$$

注记 2 在定理3.2中, 若取$Y_s=\mathbf{1}_n$, $s=1,2,\ldots,K$, 并记$\beta_0:=\mu_0$, $M=a$, $N=\sigma^2$, 则$V_i=I_n$,

$$\Delta_s = \mathbf{1}_n'(\sigma_0^2 I_n + \mathbf{1}_n\sigma^2\mathbf{1}_n')^{-1}\mathbf{1}_n = \frac{n}{\sigma_0^2+n\sigma^2}, \qquad \Delta = \frac{nK}{\sigma_0^2+n\sigma^2}, \tag{3.20}$$

以及

$$\widehat{\beta}_{\mathrm{GL}(s)} = \overline{X}_s, \qquad \widehat{\beta}_{\mathrm{GL}} = \overline{\overline{X}}. \tag{3.21}$$

因此, 信度因子为

$$Z_{i1} = \frac{n\sigma^2}{\sigma_0^2+n\sigma^2}, \qquad Z_{i2} = \frac{nKa\sigma_0^2}{(\sigma_0^2+n\sigma^2)(\sigma_0^2+n\sigma^2+nKa)}$$

以及

$$1 - Z_{i1} - Z_{i2} = \frac{\sigma_0^2}{\sigma_0^2+n\sigma^2+nKa}.$$

则根据定理3.2, 风险参数$\mu(\Theta_i,\Lambda)$的信度估计为

$$\widehat{\mu(\Theta_i,\Lambda)}^* = \frac{n\sigma^2}{\sigma_0^2+n\sigma^2}\overline{X}_i + \frac{nKa\sigma_0^2}{(\sigma_0^2+n\sigma^2)(\sigma_0^2+n\sigma^2+nKa)}\overline{\overline{X}} + \frac{\sigma_0^2}{\sigma_0^2+n\sigma^2+nKa}\mu_0. \tag{3.22}$$

式(3.22)正是具有共同效应的Bühlmann信度估计, 可参考[6]或[7].

注记 3 在定理3.2中, 若$M=0$, 即风险之间不存在风险相依, 这时$Z_{i2}=0$, 则此时信度保费为

$$\widehat{Y_i\beta(\Theta_i)}^* = Y_i(Z_{i1}\widehat{\beta}_{\mathrm{GL}(i)} + (I_p - Z_{i1})\beta_0). \tag{3.23}$$

式(3.23)正是Hachemeister回归信度保费估计.

§4. 结 论

在经典的信度回归模型中, 数据的时间分量上假设为不独立, 并且具有已知协差效应. 但一般假设保单风险之间为相互独立, 我们在[6]和[7]的模型框架下, 研究了风险之间存在共同效应导致的相依情形, 得到保险合同的回归信度保费与风险参数的回归信度估计. 我们将会进一步研究更一般的风险相依结构下的信度模型的定价问题; 其次, 本文中风险之间的相依结构(这里表现为协方差矩阵)假定为已知, 如果相依结构未知(这是实务中更为常见的情况), 相关的信度问题的估计和推断也是一个有着实际意义的问题. 最后, 如果使用的不是净保费原理(相当于平方损失), 相应的信度问题也是我们今后的研究对象之一.

参 考 文 献

[1] Hachemeister, C.A., Credibility for regression models with application to trend, in *Credibility, Theory and Application*, Proveedings of the Berkeley Acturial Research Conference on Credibility, Academic Press, New York, 1975, 129–163.

[2] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D., The concept of comonotonicity in actuarial science and finance: theory, *Insurance: Mathematics and Economics*, **31**(2002), 3–33.

[3] Dhaene, J. and Goovaerts, M.J., Dependency of risks and stop-loss order, *Astin Bulletin*, **26**(1996), 201–212.

[4] Goovaerts, M.J. and Dhaene, J., The compound Poisson approximation for a portfolio of dependent risks, *Insurance: Mathematics and Economics*, **18(1)**(1996), 81–85.

[5] Müller, A., Stop-loss order for portfolios of dependent risks, *Insurance: Mathematics and Economics*, **21**(1997), 219–223.

[6] Yeo, K.L. and Valdez, E.A., Claim dependence with common effects in credibility models, *Insurance: Mathematics and Economics*, **38**(2006), 609–629.

[7] Wen. L., Wu. X. and Zhou., X., The credibility premiums for models with dependence induced by common effects, *Insurance: Mathematics and Economics*, **44**(2009), 19–25.

[8] Rao, R. and Toutenburg, H., *Linear Models*, Springer, New York, 1995.

[9] Bühlmann, H., Experience rating and credibility, *Astin Bulletin*, **4**(1967), 199–207.

[10] Bühlmann, H. and Gisler, A., *A Course in Credibility Theory and Its Applications*, Springer, Netherlands, 2005.

The Regression Credibility Models with Random Common Effects

WANG ZHUJUAN

(*Shanghai Institute of Technology, Shanghai, 200235*)

WEN LIMIN

(*Jiangxi Normal University, Nanchang, 330022*)

In classical regression credibility models suggested by Hachemeister (1975), the risks are assumed to be mutually independent. In this paper, we introduce a dependence between risks induced by common effects and developed a credibility regression model with dependence and the credibility predictors of future claims and the estimators of risk parameters are derived under this model. The results show that the credibility estimators remain the weighted sums of individual and collective premium.

Keywords: Credibility estimator, random common effects, orthogonal projection, regression credibility.

AMS Subject Classification: 62G35.