

Quantile Credibility Models under Common Effect Structure

Ji Cheng*, Lijun Wu#

College of Mathematics and System Sciences, Xinjiang University, Urumqi Xinjiang
Email: 1960328913@qq.com, #15099376827@163.com

Received: Mar. 19th, 2019; accepted: Apr. 3rd, 2019; published: Apr. 10th, 2019

Abstract

In non-life insurance application research, the classical Bühlmann credibility model cannot effectively reflect the tail information of distribution, and the policy between the risks is not completely independent. Therefore, considering the common effect between risks and combining the credibility theory from the perspective of quantile, the quantile credibility models under common effect structure are studied, which further expands the classical credibility model.

Keywords

Common Effect, Orthogonal Projection, Quantile, Credibility Model

共同效应结构下的分位数信度模型

程 纪*, 吴黎军#

新疆大学, 数学与系统科学学院, 新疆 乌鲁木齐
Email: 1960328913@qq.com, #15099376827@163.com

收稿日期: 2019年3月19日; 录用日期: 2019年4月3日; 发布日期: 2019年4月10日

摘 要

在非寿险应用研究中, 经典的Bühlmann信度模型并不能充分利用数据尾部的信息, 且保单之间的风险并不是完全独立的, 因而考虑在分位数视角下结合信度理论, 综合风险之间存在的共同效应, 研究了具有共同效应的分位数信度模型, 拓展了经典的信度模型。

*第一作者。

#通讯作者。

关键词

共同效应, 正交投影, 分位数, 信度模型

Copyright © 2019 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在保险实践研究中, 针对特定的投保人, 观察到 n 年的索赔数据 X_1, X_2, \dots, X_n , 假设在同一费率级别中每个投保人的风险水平可用一个风险参数 Θ 来表示, 不同的投保人对应不同的 Θ , 且 Θ 的分布为 $\pi(\Theta)$, 以此将风险差异进行量化。我们假设 Θ 的存在性, 理想情形下第 $n+1$ 年应当收取个体保费(条件期望) $\mu_{n+1}(\Theta) = E(X_{n+1}|\Theta)$, 而不是净保费 $\mu_{n+1} = E(X_{n+1})$ 。

Bühlmann (1967) [1] 提出无分布信度模型, 用经验数据的线性组合 $\alpha_0 + \sum_{i=1}^n \alpha_i X_i$ 来逼近 $\mu_{n+1}(\theta)$, 通过求解下面的期望损失函数:

$$\min_{\alpha_0, \alpha_i} E \left[\mu_{n+1}(\theta) - \alpha_0 - \sum_{i=1}^n \alpha_i X_i \right]^2$$

得到最优信度保费。之后, Bühlmann 和 Straub (1970) [2] 将无分布信度模型延伸到自然权重的情形。但这些模型均要求所研究的保单数据独立同分布, 同时个体风险之间相互独立, 我们期待能够找到一种针对特定环境下对保单进行分组并计算每个保单第 $n+1$ 年的个体保费 $\mu_{n+1}(\theta)$ 的费率厘定方法。因此, Yeo and Valdez (2006) [3] 通过用随机变量刻画风险之间存在的某种共同效应(例如某种特定环境), 在这种特殊结构下研究了信度保费估计, 而后温利民等分别研究了信度模型及回归信度模型在共同效应情况下的信度保费[4] [5]。更一般地, 考虑保单风险之间的相依关系, 打破经典的风险独立假设, 张强等(2018) [6] 研究了平衡损失函数下具有相依结构的信度估计。

鉴于保险数据经常具有较强的鲁棒性, 存在一些具有强影响的离群点, 经典的 Bühlmann 信度模型并不能有效利用这些离群值的数据信息, 而分位数方法能够很好地解决这一问题, Georgios Pitselis (2013) [7] 在风险独立性的假设下, 用分位数技术首次建立了分位数信度模型, 阐述了分位数与经典的 Bühlmann 信度与 Hachemeister 回归信度模型之间的联系, 并将分位数纳入到这两大经典的信度模型中。本文综合考虑数据鲁棒性及风险之间的特定影响效应, 创造性地建立了具有共同效应的分位数信度模型, 致力于充分利用所有数据的信息, 作出最佳的估计, 从而实现保险产品结构的优化。

2. 模型假设与准备

2.1. 准备知识

考虑 K 份保单合同, 对于第 i 份合同有 n 年的索赔数据, $i=1, 2, \dots, k$, 则样本数据构成一个矩阵:

$$\begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{pmatrix}$$

风险由风险参数 $\theta \in \Theta$ 进行描述, 对任意 $p \in (0,1)$, ξ 为概率空间 (Ω, Γ, P) 上的随机变量, 令 $\xi(p) = F^{-1}(p) = \inf\{x: F(x) \geq p\}$ 为 ξ 的 p -分位数。

在给定风险参数 Θ 下, 第 i 份保单的索赔随机向量为 X_i , $X_i = (X_{i1}, X_{i2}, \dots, X_{in})'$, 设 $F_n(x)$ 为样本 $\{X_{i1}, X_{i2}, \dots, X_{in}\}$ 的经验分布函数, 则样本 P -分位数的估计可定义为:

$$\hat{\xi}(p) = n \left(\frac{j}{n} - p \right) X_{(j-1)} + n \left(p - \frac{j-1}{n} \right) X_{(j)}, \text{ 其中 } \frac{j-1}{n} \leq p \leq \frac{j}{n}, j=1, \dots, n.$$

记 $X = (X'_1, X'_2, \dots, X'_k)'$ 为所有索赔数据组成的列向量, $\hat{\xi}_{ij}(p)$ 为样本分位数的估计。我们的目标是通过样本选取不同分位点, 预测每份保单在各分位点第 $n+1$ 年的索赔。由于各保单之间存在某一特定效应, 我们用随机变量 Λ 来刻画此种共同效应, 下面给出模型的基本假设。

2.2. 模型假设

假设 1: 风险之间存在共同效应 Λ , 且 $E[\Lambda] = \mu_\lambda, \text{Var}[\Lambda] = S_\lambda^2$ 。

假设 2: 给定共同效应 Λ , 风险参数 Θ_i , 索赔样本数据为 $X_{i1}, X_{i2}, \dots, X_{in}$, 其样本分位数 $\hat{\xi}_{ij}(p)$ 的条件期望和条件方差如下:

$$E\left(\hat{\xi}_{ij}(p) | \Theta_i, \Lambda\right) = \Xi_p(\Theta_i, \Lambda), \text{Var}\left(\hat{\xi}_{ij}(p) | \Theta_i, \Lambda\right) = V_p(\Theta_i, \Lambda).$$

引入下列记号:

$$\begin{aligned} E[\Xi_p(\Theta_i, \Lambda) | \Lambda] &= \Xi_1(p, \Lambda), \text{Var}[\Xi_p(\Theta_i, \Lambda) | \Lambda] = S_2^2(p, \Lambda), \\ E[\Xi_1(p, \Lambda)] &= \Xi_1(p), E[S_2^2(p, \Lambda)] = S_2^2(p), \text{Var}[\Xi_1(p, \Lambda)] = \psi_1(p), \\ E[V_p(\Theta_i, \Lambda) | \Lambda] &= S_1^2(p, \Lambda), E[S_1^2(p, \Lambda)] = S_1^2(p). \end{aligned}$$

2.3. 模型求解

根据信度理论, 我们用样本分位数的线性函数逼近最优信度保费, 即求解最小化问题:

$$\min_{A, B} E \left[\left(\Xi(p) - A - B \hat{\xi}(p) \right) \left(\Xi(p) - A - B \hat{\xi}(p) \right)' \right] \quad (1)$$

式中,

$$\begin{aligned} \Xi(p) &= \left[\hat{\xi}_{1, n+1}(p), \dots, \hat{\xi}_{K, n+1}(p) \right]' \\ \hat{\xi}(p) &= \left[\hat{\xi}_1(p)', \dots, \hat{\xi}_K(p)' \right]' \\ \hat{\xi}_i(p) &= \left[\hat{\xi}_{i1}(p), \dots, \hat{\xi}_{in}(p) \right]' \end{aligned}$$

基于样本分位数, $\Xi(p)$ 的非齐次信度估计为 $\hat{\Xi}(p) = \text{pro} \left[\Xi(p) | L(\hat{\xi}(p), 1) \right]$, 本文利用正交投影法求解最优化问题(1), 引入下列引理。

引理 1: 设 $X = (X_1, \dots, X_n)'$, $Y = (Y_1, \dots, Y_m)'$ 为两个随机向量, Y 在 $L(X, 1)$ 上的正交投影为

$$\text{pro}(Y | L(X, 1)) = E(Y) + \text{Cov}(Y, X) \text{Cov}^{-1}(X, X) (X - E(X))$$

证明见 Wen *et al.* (2009) [4]。

引理 2: 设 A, B, C, D 为适当阶数的矩阵, 且 A, C 可逆, 则有求逆公式:

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

证明: 由

$$\begin{aligned} & (A + BCD) \left[A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \right] \\ &= I - B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} + BCDA^{-1} - BCDA^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \\ &= I + BCDA^{-1} - B(I + CDA^{-1}B)(C^{-1} + DA^{-1}B)^{-1}C^{-1}CDA^{-1} \\ &= I + BCDA^{-1} - B(I + CDA^{-1}B)(I + CDA^{-1}B)^{-1}CDA^{-1} \\ &= I + BCDA^{-1} - BCDA^{-1} \\ &= I \end{aligned}$$

即

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

引理 3: 在模型假设 1 和假设 2 下, 有以下结果:

1) $\hat{\xi}_i(p) = [\hat{\xi}_{i1}(p), \dots, \hat{\xi}_{in}(p)]'$ 的期望为:

$$E[\hat{\xi}_i(p)] = \Xi_i(p)1_n.$$

其中 1_n 为所有元素均为 1 的 n 维列向量。

2) 向量 $\hat{\xi}(p) = [\hat{\xi}_1(p)', \dots, \hat{\xi}_k(p)']'$ 的协方差矩阵为:

$$\text{Cov}[\hat{\xi}(p), \hat{\xi}(p)] = I_k \otimes (S_1^2(p)I_n + S_2^2(p)1_n1_n') + \psi_1(p)1_{nk}1_{nk}'.$$

3) 矩阵 $\text{Cov}[\hat{\xi}(p), \hat{\xi}(p)]$ 的逆矩阵为:

$$\begin{aligned} \text{Cov}^{-1}[\hat{\xi}(p), \hat{\xi}(p)] &= \frac{1}{S_1^2(p)} I_k \otimes \left(I_n - \frac{S_2^2(p)}{S_1^2(p) + nS_2^2(p)} 1_n1_n' \right) \\ &\quad - \frac{\psi_1(p)}{[S_1^2(p) + nS_2^2(p)][S_1^2(p) + nS_2^2(p) + nk\psi_1(p)]} 1_{nk}1_{nk}'. \end{aligned}$$

4) $\Xi(p) = [\hat{\xi}_{1,n+1}(p), \dots, \hat{\xi}_{k,n+1}(p)]'$ 与 $\hat{\xi}(p) = [\hat{\xi}_1(p)', \dots, \hat{\xi}_k(p)']'$ 的协方差矩阵为

$$\text{Cov}[\Xi(p), \hat{\xi}(p)] = (S_2^2(p)e_i' + \psi_1(p)1_k') \otimes 1_n'.$$

证明:

1)

$$E[\hat{\xi}_i(p)] = E[E[\hat{\xi}_i(p)|\Theta_i, \Lambda]] = E \begin{pmatrix} E(\Xi_p(\Theta_i, \Lambda)|\Lambda) \\ \vdots \\ E(\Xi_p(\Theta_i, \Lambda)|\Lambda) \end{pmatrix} = E \begin{pmatrix} \Xi_i(p, \Lambda) \\ \vdots \\ \Xi_i(p, \Lambda) \end{pmatrix} = \Xi_i(p)1_n.$$

2) 记 $\Theta = (\Theta_1, \dots, \Theta_k)'$, 则

$$\begin{aligned} E[\hat{\xi}(p)|\Theta, \Lambda] &= [\Xi_p(\Theta_1, \Lambda)1_n', \dots, \Xi_p(\Theta_k, \Lambda)1_n']' = [\Xi_p(\Theta_1, \Lambda), \dots, \Xi_p(\Theta_k, \Lambda)]' \otimes 1_n \\ \text{Cov}[\hat{\xi}(p)|\Theta, \Lambda] &= \text{diag}[V_p(\Theta_1, \Lambda)I_n, \dots, V_p(\Theta_k, \Lambda)I_n] \\ \text{Cov}[\hat{\xi}(p), \hat{\xi}(p)] &= \text{Cov}[E(\hat{\xi}(p)|\Theta, \Lambda)] + E[\text{Cov}(\hat{\xi}(p)|\Theta, \Lambda)] \triangleq I_1 + I_2 \end{aligned}$$

分项计算得:

$$\begin{aligned}
 I_1 &= \text{Cov}\left[E\left(\hat{\xi}(p)\right)\left|\Theta, \Lambda\right.\right] \\
 &= E\left(\text{Cov}\left\{\left[\Xi_p\left(\Theta_1, \Lambda\right), \dots, \Xi_p\left(\Theta_k, \Lambda\right)\right] \otimes 1_n \mid \Lambda\right\}\right) + \text{Cov}\left(E\left\{\left[\Xi_p\left(\Theta_1, \Lambda\right), \dots, \Xi_p\left(\Theta_k, \Lambda\right)\right] \otimes 1_n \mid \Lambda\right\}\right) \\
 &= E\left[S_2^2(p, \Lambda)\right] I_k \otimes 1_n 1_n' + \text{Cov}\left[\Xi_1(p, \Lambda) 1_{nk}\right] \\
 &= S_2^2(p) I_k \otimes 1_n 1_n' + \psi_1(p) 1_{nk} 1_{nk}' \\
 I_2 &= E\left[\text{Cov}\left(\hat{\xi}(p)\right)\left|\Theta, \Lambda\right.\right] = S_1^2(p) I_{nk}
 \end{aligned}$$

则有,

$$\text{Cov}\left[\hat{\xi}(p), \hat{\xi}(p)\right] \triangleq I_1 + I_2 = I_k \otimes \left(S_1^2(p) I_n + S_2^2(p) 1_n 1_n'\right) + \psi_1(p) 1_{nk} 1_{nk}'$$

3) 根据引理 2 得

$$\begin{aligned}
 \left(S_1^2(p) I_n + S_2^2(p) 1_n 1_n'\right)^{-1} &= \frac{1}{S_1^2(p)} I_n - \frac{(p) 1_n 1_n'}{\frac{1}{S_2^2(p)} + \frac{1}{S_1^2(p)} 1_n 1_n'} \\
 &= \frac{1}{S_1^2(p)} \left(I_n - \frac{S_2^2(p)}{S_1^2(p) + n S_2^2(p)} 1_n 1_n'\right)
 \end{aligned}$$

$$\text{因此, } \left(S_1^2(p) I_n + S_2^2(p) 1_n 1_n'\right)^{-1} 1_n = \frac{1}{S_1^2(p) + n S_2^2(p)} 1_n.$$

因为 $1_{nk} = 1_k \otimes 1_n$.

所以,

$$\begin{aligned}
 \text{Cov}^{-1}\left[\hat{\xi}(p), \hat{\xi}(p)\right] &= I_k \otimes \left(S_1^2(p) I_n + S_2^2(p) 1_n 1_n'\right)^{-1} \\
 &\quad \frac{\left[I_k \otimes \left(S_1^2(p) I_n + S_2^2(p) 1_n 1_n'\right)^{-1}\right] 1_{nk} 1_{nk}' \left[I_k \otimes \left(S_1^2(p) I_n + S_2^2(p) 1_n 1_n'\right)^{-1}\right]}{\frac{1}{\psi_1(p)} + 1_{nk}' \left[I_k \otimes \left(S_1^2(p) I_n + S_2^2(p) 1_n 1_n'\right)^{-1}\right] 1_{nk}} \\
 &= \frac{1}{S_1^2(p)} I_k \otimes \left(I_n - \frac{S_2^2(p)}{S_1^2(p) + n S_2^2(p)} 1_n 1_n'\right) \\
 &\quad \frac{\psi_1(p)}{\left[S_1^2(p) + n S_2^2(p)\right] \left[S_1^2(p) + n S_2^2(p) + nk \psi_1(p)\right]} 1_{nk} 1_{nk}'.
 \end{aligned}$$

4) 因为 $\text{Cov}\left(\Xi(p), \hat{\xi}(p)\right)\left|\Theta, \Lambda\right. = 0$, 则

$$\begin{aligned}
 \text{Cov}\left[\Xi(p), \hat{\xi}(p)\right] &= \text{Cov}\left[E\left(\Xi(p)\right)\left|\Theta, \Lambda\right.\right], E\left(\hat{\xi}(p)\right)\left|\Theta, \Lambda\right.\right] + E\left[\text{Cov}\left(\Xi(p), \hat{\xi}(p)\right)\left|\Theta, \Lambda\right.\right] \\
 &= \text{Cov}\left\{\Xi_p\left(\Theta_i, \Lambda\right), \left[\Xi_p\left(\Theta_1, \Lambda\right), \dots, \Xi_p\left(\Theta_k, \Lambda\right)\right] \otimes 1_n\right\} \\
 &= E\left(\text{Cov}\left\{\Xi_p\left(\Theta_i, \Lambda\right), \left[\Xi_p\left(\Theta_1, \Lambda\right), \dots, \Xi_p\left(\Theta_k, \Lambda\right)\right] \otimes 1_n \mid \Lambda\right\}\right) \\
 &\quad + \text{Cov}\left(E\left[\Xi_p\left(\Theta_i, \Lambda\right)\right], \left\{E\left[\Xi_p\left(\Theta_1, \Lambda\right)\right], \dots, E\left[\Xi_p\left(\Theta_k, \Lambda\right)\right]\right\} \otimes 1_n\right) \\
 &= S_2^2(p) e_i' \otimes 1_n' + \text{Cov}\left[\Xi_1(p, \Lambda), \Xi_1(p, \Lambda) 1_{nk}\right] = S_2^2(p) e_i' \otimes 1_n' + \psi_1(p) 1_{nk}' \\
 &= \left(S_2^2(p) e_i' + \psi_1(p) 1_k'\right) \otimes 1_n'
 \end{aligned}$$

引理 3 证毕。

3. 具有共同效应结构的分位数信度模型

定理 3.1: 在模型假设 1 及假设 2 下, 求解最小化问题(1)可得 $\Xi_p(\Theta)$ 的最优线性信度估计为

$$\hat{\Xi}(p) = Z_1 \bar{\xi}(p) + Z_2 \bar{\bar{\xi}}(p) + Z_3 \Xi_1(p)$$

式中,

$$Z_1 = \frac{nS_2^2(p)}{nS_2^2(p) + S_1^2(p)},$$

$$Z_2 = \frac{nk\psi_1(p)S_1^2(p)}{[nS_2^2(p) + S_1^2(p)][nk\psi_1(p) + nS_2^2(p) + S_1^2(p)]},$$

$$Z_3 = 1 - Z_1 - Z_2 = \frac{S_1^2(p)}{nk\psi_1(p) + nS_2^2(p) + S_1^2(p)}.$$

为信度因子。

其中,

$$\bar{\xi}_i(p) = \frac{1}{n} \sum_{j=1}^n \hat{\xi}_{ij}(p), \bar{\bar{\xi}}(p) = \frac{1}{k} \sum_{i=1}^k \bar{\xi}_i(p).$$

证明: 由引理 1 及引理 3 得

$$\begin{aligned} \hat{\Xi}(p) &= E[\Xi(p)] + Cov[\Xi(p), \hat{\xi}(p)] Cov^{-1}[\hat{\xi}(p), \hat{\xi}(p)] [\hat{\xi}(p) - E(\hat{\xi}(p))] \\ &= \Xi_1(p) + \frac{1}{nS_2^2(p) + S_1^2(p)} \left[S_2^2(p) e'_i \otimes 1'_n + \frac{\psi_1(p) S_1^2(p)}{nk\psi_1(p) + nS_2^2(p) + S_1^2(p)} 1'_{nk} \right] [\hat{\xi}(p) - \Xi_1(p, \Lambda) 1_{nk}] \\ &= \Xi_1(p) + \frac{1}{nS_2^2(p) + S_1^2(p)} \left[nS_2^2(p) (\bar{\xi}_i(p) - \Xi_1(p)) + \frac{nk\psi_1(p) S_1^2(p)}{nk\psi_1(p) + nS_2^2(p) + S_1^2(p)} (\bar{\bar{\xi}}(p) - \Xi_1(p)) \right] \\ &= Z_1 \bar{\xi}_i(p) + Z_2 \bar{\bar{\xi}}(p) + Z_3 \Xi_1(p) \end{aligned}$$

定理 3.1 得证。

4. 总结

本文在分位数的视角下, 结合信度理论, 建立了共同效应结构下的分位数信度模型。通过自由选择样本分位点打破了经典的信度模型中数据鲁棒性的弊端, 利用随机变量反映风险之间的共同效应, 解决实际应用中保单风险并不独立的情形, 得到了最优的信度保费, 从而降低经营风险, 使得保险产品真正回归到保障本质。

致 谢

感谢国家自然科学基金课题“基于分位数回归的信度理论与风险度量的研究”(11861064)的资助, 为本研究工作提供了良好的平台; 感谢课题组导师在课题研究中无微不至地传道、授业、解惑, 使我不断进步; 感谢课题组成员对我的帮助、关心, 伴我一路成长。

参考文献

- [1] Bühlmann, H. (1967) Experience Rating and Credibility. *ASTIN Bulletin*, **5**, 199-207.

<https://doi.org/10.1017/S0515036100008989>

- [2] Bühlmann, H. and Straub, E. (1970) Glaubwürdigkeit für Schadensätze. *Bulletin of the Swiss Association of Actuaries*, **70**, 111-133.
- [3] Yeo, K.L. and Valdez, E.A. (2006) Claim Dependence with Common Effects in Credibility Models. *Insurance Mathematics & Economics*, **38**, 609-629. <https://doi.org/10.1016/j.insmatheco.2005.12.006>
- [4] Wen, L., Wu, X. and Zhou, X. (2009) The Credibility Premiums for Models with Dependence Induced by Common Effects. *Insurance Mathematics & Economics*, **44**, 19-25. <https://doi.org/10.1016/j.insmatheco.2008.09.005>
- [5] Wang, Z. and Wen, L. (2011) Regression Credibility Models with Random Common Effects. *Chinese Journal of Applied Probability & Statistics*, **27**, 312-322.
- [6] Zhang, Q. and Chen, P. (2018) Credibility Estimators with Dependence Structure over Risks and Time under Balanced Loss Function. *Statistica Neerlandica*, **72**, 153-179.
- [7] Pitselis, G. (2013) Quantile Credibility Models. *Insurance Mathematics & Economics*, **52**, 477-489. <https://doi.org/10.1016/j.insmatheco.2013.02.011>

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: sa@hanspub.org