Dependence Analysis of Insurance Businesses Based on Hierarchical Archimedean Copula Function

Jing Tian
School of Mathematics, Tianjin University, Tianjin 300350, China.
tianjing95@tju.edu.cn

Abstract. The dependence between insurance businesses greatly influences the evaluation of insurance companies' overall reserves. Because of the heterogeneity of risks, the dependence between businesses is often not exactly the same. In order to analyze the asymmetric dependence relationship, an asymmetric dependence structure based on hierarchical Archimedean Copula function is proposed. It is assumed that the loss distribution of business lines follows Log-normal, Gamma and Weibull distribution. And then a hierarchical Copula dependence structure among different businesses is established. By the empirical analysis of the historical data of CPIC's automobile insurance, transportation insurance, enterprise property insurance and liability insurance, the optimal 4-dim hierarchical Copula structure is selected to analyze the dependence among the four. The empirical results show that automobile insurance and transportation insurance have the strongest dependence, followed by enterprise property insurance and liability insurance. At the same time, based on the obtained hierarchical structure, reserve assessment and risk measurement can be carried out, which provides an insight for analyzing the dependence among multivariate analysis.

Keywords: Dependence analysis; hierarchical archimedean copula; loss distribution; multivariate analysis.

1. Introduction
The dependence between different business lines of an insurance company greatly influences the evaluation of reserves. Generally, the different businesses of insurance companies are neither completely independent nor completely related. And there always exists a nonlinear dependence relationship. For example, in road traffic accidents, automobile insurance and accident liability insurance tend to be non-linearly related. So, it is very important to adopt appropriate methods to determine the reasonable dependence between business lines for the evaluation of insurance companies' reserves.

At present, Copula function is widely used in dependence analysis. It allows the separation of dependence structures from marginal distributions, making it a flexible and powerful analysis tool for dependent data. It can analyze not only linear correlation but also nonlinear correlation, so it is popular among scholars [1-4]. Nowadays, elliptic Copula and Archimedean Copula are mostly used in existing literature. However, the hierarchical Archimedean Copula [5-7] not only can avoid the higher dimensional integration of elliptic Copula in deriving joint probability but also overcome that...
Archimedean Copula can only describe symmetric dependence. Therefore, we propose an asymmetric dependence structure based on hierarchical Archimedean Copula. Because of the asymmetric, non-negative domain and thick tails of marginal distribution, we choose Log-normal, Gamma and Weibull distributions in the fitting of margins [8].

2. The Dependence Model

2.1. Hierarchical Archimedean Copula

The basic idea of HAC model is to construct the hierarchical structure of Archimedean Copulas. The dependence between the variables decreases as the HAC level increases, and the Copula parameter describing the dependence decreases as well. There are two kinds of HAC: partially nested and fully nested. As Figure 1 shows, A is fully nested and B, C are partially nested. A 4-dimensional fully nested HAC is defined as:

\[ C(u_1, u_2, u_3, u_4) = \phi_3^{-1}(\phi_1^{-1}(\phi_2^{-1}(\phi_1(u_1) + \phi_2(u_2)) + \phi_3(u_3)) + \phi_4(u_4)) \]  

(1)

Where \( \phi_1, \phi_2, \phi_3 \) are generator functions. The widely used generators are Clayton, Gumbel and Frank.

2.2. Marginal Distribution

Accounting for the sharp peak and thick tail of the claim data, we adopt Gamma distribution, Log-normal distribution and Weibull distribution for the margins.

(1) The density of Gamma distribution[8]:

\[ f(x|\alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \]  

(2)

where \( \alpha \) is the shape parameter, \( \beta \) is the scale parameter.

(2) The density of Log-normal distribution[8]:

\[ f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi x}\sigma} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \]  

(3)

where, \( \mu \) is the mean of the log of variables, \( \sigma \) is the standard deviation of the log of variables.

(3) The density of Weibull distribution[8]:

\[ f(x|\alpha, \beta) = \frac{\beta}{\alpha} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}} \]  

(4)

where \( \beta, \alpha \) are shape and scale parameter respectively.

Fig. 1 Network structure of different HAC

According to Savu and Trede[9], a hierarchical model must satisfy: (a) the higher the level, the smaller the parameters, as shown in Figure 1; (b) for \( i = 1, 2, \ldots, n \), \( \phi_i^{-1} \) is completely monotonous; (c) suppose \( \omega = \phi_i^{-1} \circ \phi_i \) satisfy \( \omega \in L^\infty \), where

\[ L = \{ \omega : [0, \infty) \rightarrow [0, \infty) | \omega(0) = 0, \omega(\infty) = \infty, (-1)^{i-1} \omega^{(i)} \geq 0, j = 1, \ldots, n \} \]  

where, \( \omega \) is a function.

2.2. Marginal Distribution

Accounting for the sharp peak and thick tail of the claim data, we adopt Gamma distribution, Log-normal distribution and Weibull distribution for the margins.

(1) The density of Gamma distribution[8]:

\[ f(x|\alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \]  

(2)

where \( \alpha \) is the shape parameter, \( \beta \) is the scale parameter.

(2) The density of Log-normal distribution[8]:

\[ f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi x}\sigma} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \]  

(3)

where, \( \mu \) is the mean of the log of variables, \( \sigma \) is the standard deviation of the log of variables.

(3) The density of Weibull distribution[8]:

\[ f(x|\alpha, \beta) = \frac{\beta}{\alpha} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}} \]  

(4)

where \( \beta, \alpha \) are shape and scale parameter respectively.
\[ f(x|\lambda, k) = \frac{k}{\lambda} \left( \frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k} ; x \geq 0 \] (4)

where \( \lambda \) is the shape parameter, \( k \) is the scale parameter.

3. Real Data Application

3.1. The Data

We use the historical claims data of China Pacific Property Insurance Company (CPIC) (enterprise property insurance (EPI), automobile insurance (AMI), transportation insurance (TI) and liability insurance (LI)) to analyze the dependence between insurance business lines. Although the business setting of CPIC changed slightly from 1997 to 2017, the four lines were always established and accounted for as much as 90% of the premium income of CPIC over the years. Our data comes from the claim data of the four businesses of CPIC from 1997 to 2017, and the data are extracted from China Insurance Yearbook.

Figure 2 shows the dependence matrix analysis of the four lines. As can be seen from the diagonal of the figure, these data are asymmetric and right-skewed, indicating that the whole sequence is more likely to be non-normal distributed. At the same time, from the triangular part of Figure 2, there is a significant correlation between the four business lines.

![Fig. 2 Dependence matrix analysis of 4 business lines](image)

3.2. Estimation of Margins

The total claims of the insurance company mainly depends on the tails of the theoretical distribution of businesses and the dependence among them. In this paper, 3 distributions commonly used in actuarial theory are selected to fit the data: Log-normal, Gamma and Weibull distribution. All three distributions are determined by two parameters, but have tails of different thicknesses. The parameters of the model are estimated by the maximum likelihood method. The optimal results are shown in Table 1. The Q-Q plot of the fitted distribution and theoretical distribution (see Fig. 3) also shows that the marginal distribution has a good fitting.
Table 1. Fitting results of marginal distribution

| Distribution | Parameter 1 | Parameter 2 | AIC     |
|--------------|-------------|-------------|---------|
| EPI Gamma    | shape=2.19  | rate=1.30e-3| -174.40 |
| AMI Log-normal | log(mean)=9.15 | log(sd)=1.15 | -224.76 |
| TI Gamma     | shape=2.16  | rate=4.60e-3| -147.28 |
| LI Log-normal | log(mean)=5.13 | log(sd)=2.25 | -154.56 |

Fig. 3 Q-Q plot of the marginal distributions

3.3. Estimation of Dependence Structure

According to the above marginal distribution, the probability integral transformation is carried out to obtain the uniform distribution on [0,1]. Three widely used generators, Gumbel, Clayton and Frank, are selected to construct HAC. According to Akaike Information Criterion (AIC), the smaller the AIC value, the better the fitting. By comparing the AIC values of the model in Table 2, it can be seen that the optimal fitting is 4-dim Gumbel-HAC.

At the same time, draw the plot of the empirical Copula and the fitted Copula. As can be seen from Figure 4, Gumbel function has the optimal fitting. The results obtained are the same as the numerical results. Thus, among the four business lines, automobile insurance and transportation insurance have the strongest correlation, followed by enterprise property insurance and finally liability insurance.

Table 2. Estimation results of 4-dim HAC

|          | C1-Gumbel | C2-Clayton | C3-Frank |
|----------|-----------|------------|----------|
| $\theta_1$ | $\hat{\theta}_{(AMI, TI)} = 8.05$ | $\hat{\theta}_{(EPI, AMI)} = 7.15$ | $\hat{\theta}_{(AMI, TI)} = 29.46$ |
| $\theta_2$ | $\hat{\theta}_{(AMLTI,EPI)} = 5.82$ | $\hat{\theta}_{(EPI,AMI, TI)} = 5.47$ | $\hat{\theta}_{(AMLTI,EPI)} = 19.56$ |
| $\theta_3$ | $\hat{\theta}_{(AMLTI,EPI, LI)} = 3.03$ | $\hat{\theta}_{(EPLAMI, TI, LI)} = 1.12$ | $\hat{\theta}_{(AMLTI,EPI, LI)} = 8.88$ |

AIC: -154.55 -108.42 -128.21
4. Summary
Based on the hierarchical Archimedean Copula, the asymmetric dependence structure among automobile insurance, transportation insurance, enterprise property insurance and liability insurance is established. By comparing the fitted results with the empirical model, the optimal hierarchical model is the 4-dim Gumbel hierarchical structure. The empirical results show that the dependence of the four insurance businesses is different. Automobile insurance and transportation insurance have the strongest dependence, followed by enterprise property insurance and finally liability insurance. Therefore, when evaluating the overall reserves, insurance companies should fully consider the asymmetric dependence among business lines, which is helpful to obtain more reasonable reserves. This paper provides an insight for risk management and the dependence analysis of multivariate analysis.

References
[1] Nelsen R B. An Introduction to Copulas. Springer, 2006, p. 157-225.
[2] Giampiero M, Rosalba R. Copula Link-Based Additive Models for Right-Censored Event Time Data. Journal of the American Statistical Association. Vol. 115 (2020) No. 530, p. 886-895.
[3] Shi P, Zhang W. A copula regression model for estimating firm efficiency in the insurance industry. Journal of Applied Statistics. Vol. 38 (2011) No. 10, p. 2271-2287.
[4] Yang L, Shi P. Multi-peril rate making for property insurance using longitudinal data. Journal of the Royal Statistical Society Series A-Statistics in Society. Vol. 182 (2019) No. 2, p. 647-668.
[5] Joe H. Dependence Modeling with Copulas. CRC Press, 2015, p. 9-97.
[6] Deng Y, Chaganty N R. Hierarchical Archimedean copula models for the analysis of binary familial data. Statistics in Medicine. Vol. 37 (2018) No. 4, p. 590-597.
[7] Li H, Lu Y. Modeling cause-of-death mortality using hierarchical Archimedean copula. Scandinavian Actuarial Journal. Vol. 2019 (2019), No. 3, p. 247-272.
[8] Klugman S A, Panjer H H, Wilmot G E. Loss Models: From Data to Decisions. John Wiley, 2019, p. 61-80.
[9] Savu C, Trede M. Hierarchies of Archimedean copulas. Quantitative Finance. Vol. 10 (2010) No. 3, p. 295-304.