İstatistikçiler Dergisi: İstatistik & Aktüerya

Journal of Statisticians: Statistics and Actuarial Sciences

IDIA 13, 2020, 1, 18-29

Geliş/Received:02.03.2020, Kabul/Accepted: 08.06.2020

www.istatistikciler.org

Araştırma Makalesi / Research Article

A comparative study on modeling of dependence between claim severity and frequency with Archimedean copulas

Övgücan Karadağ Erdemir

Hacettepe University
Department of Actuarial Science
06800-Çankaya, Ankara, Turkey
ovgucan@hacettepe.edu.tr
0000-0002-4725-3588

Meral Sucu

Hacettepe University
Department of Actuarial Science
06800-Çankaya, Ankara, Turkey
msucu@hacettepe.edu.tr
0000-0002-7991-1792

Abstract

In the estimation of aggregate loss, claim severity and claim frequency are generally assumed independent. Although, the independence assumption is quite basic, it may cause underestimates or overestimates in calculations. The dependence in the estimation of aggregate loss can be included with the copula-based models. The joint cumulative distribution and the joint probability density functions of mixed variables such as continuous claim severity and discrete claim frequency can be obtained using the bivariate copula functions and the mixed copula approach. In this study, aggregate loss is modeled using the bivariate Archimedean copula functions considering the dependency between claim components. It is assumed that claim severity and frequency have Gamma and zero-truncated Poisson distributions, respectively. In the application part, the aggregate loss is calculated using different Archimedean copula functions. Different copula functions and different parameters for each copula are used to analyze the effect of copula type and parameter. Furthermore, aggregate loss in the presence of dependence between claims are estimated.

Anahtar sözcükler: Bivariate Archimedean copula, Copula-based model, Dependence, Mixed-variable.

Öz

Hasar tutarı ve sayısı arasındaki bağımlılığın Arşimet kopulalar ile modellenmesi üzerine karşılaştırmalı bir çalışma

Toplam hasar tahmininde, hasar tutarı ile hasar sayısı genellikle bağımsız varsayılmaktadır. Bağımsızlık varsayımı oldukça temel olmasına rağmen, hesaplamalarda gerçek değerden daha düşük veya daha yüksek tahminlere neden olabilir. Toplam hasar tahminindeki bağımlılık, kopula-temelli modeller ile dikkate alınabilir. Sürekli hasar tutarı ve kesikli hasar sayısı gibi karma değişkenlerin ortak kümülatif dağılım ve ortak olasılık yoğunluk fonksiyonu iki değişkenli kopula fonksiyonları ve karma kopula yaklaşımı kullanılarak elde edilebilir. Bu çalışmada toplam hasar, hasar bileşenleri arasındaki bağımlılık dikkate alınarak, Arşimet kopula fonksiyonları yardımıyla modellenmiştir. Hasar tutarı ve sayısının sırasıyla Gamma ve sıfır-kesilmiş Poisson dağılımlarına sahip olduğu varsayılmıştır. Uygulama bölümünde, toplam hasar farklı Arşimet kopulalar kullanılarak tahmin edilmiştir. Kopula türü ve parametresinin etkisini analiz etmek amacıyla farklı kopula fonksiyonları ve her bir kopula için farklı parametreler kullanılmıştır.

Keywords: *İki değişkenli Arşimet kopula, Kopula-temelli model, Bağımlılık, Karma-değişken.*

1. Introduction

Claim severity and frequency are the main components of non-life insurance mathematics. Analyses such as ratemaking, reserve calculation and credibility are generally carried out under the independence assumption between the claim components. According to the compound risk model, claim severity and frequency are assumed independent, and aggregate loss is calculated by Lundberg [1]. The aggregate loss is termed as risk premium corresponding only the expected loss without expenses. In ratemaking studies, generally claim severity and frequency are modeled by generalized linear models separately. Thereafter, expected values of claim severity and frequency are multiplied to calculate the aggregate loss under independence assumption.

Although, the independence assumption is quite essential, it causes underestimate or overestimate of aggregate loss compared the actual aggregate loss. To eliminate the effects of the independence assumption, some approaches have been proposed to model the dependence. Copula which is the most used method for modeling dependence in financial and statistical studies is also introduced in actuarial studies by Frees and Valdez [2]. Studies about dependence modeling using copulas can be divided into two part with and without GLMs. Song [3] defined the mixed copula approach using GLMs and Gauss copula function. The mixed copula approach lays the groundwork to model dependency between mixed variables such as continuous claim severity and discrete claim frequency. Kastenmeier [4] established a joint regression model for claim severity and frequency using the mixed copula approach. Song et al. [5], Kolev and Paiva [6] and Czado et al. [7] also modeled dependency between mixed variables. Czado et al. [7] modeled dependency between claim severity and frequency using Gauss copula and marginal GLMs. Krämer et al. [8] used Clayton, Gumbel and Frank copulas besides Gauss copula and they referred the models contain GLMs and copula as the copula-based regression models. Krämer et al. [8] also modeled the dependency between claim components only using the copula and claims without GLMs and the approach is entitled as the copula-based models. Copula-based models are useful for modeling the dependence according to only the distribution of claim components without any explanatory variables.

Besides copula-based models and copula-based regression models, Gschlößl and Czado [9] introduced a new approach to model dependency between claim severity and frequency by taking the claim frequency as an explanatory variable in the GLM modeling of aggregate loss. Garrido et al. [10] also used same approach to model dependency in non-life insurance.

Except copula; markov models [11], continuous time processes [12] and multivariate distributions such as phase-type distributions [13] are also used to model dependence in actuarial sciences.

This study is important with regars to showing that the copula function, has also started to be used in non-life insurance mathematics for dependence modeling though copula-based models. It has been also shown that the dependency between claim components can be modeled flexibly with different copula functions and different parameters.

In this study, dependency between claim components is modeled via copula-based models using Archimedean copula family. This paper is organized as follows: Firstly, bivariate Archimedean copula functions are given in Section 2. The properties of aggregate loss under the dependence and independence assumptions and the relationship between copula and aggregate loss are briefly mentioned in Section 3. An application study is carried out to compare the copulas and to estimate expected loss in the presence of dependency between claim components in Section 4 and the concluding remarks are given in Section 5.

2. Archimedean copula functions

Copula is introduced by Sklar [14] and allows us to model dependence among variables in many disciplines such as economy, finance, econometric, statistics and actuarial science. A copula is a function which joins a multivariate distribution function to its marginal distribution functions which have standart uniform distributions [15].

Archimedean copula family is one of the most frequently used parametric copulas. Archimedean copulas are diversified as Clayton, Gumbel, Frank, Ali Mikhail Haq and Joe copula functions. As Archimedean copulas are parametric copulas, they are expressed with a parameter. Using Sklar's Theorem [15], where θ is the copula parameter, a bivariate parametric copula $C(.,.|\theta)$ can be defined as follows:

$$F_{XY}(x,y|\theta) = C(F_X(x), F_Y(y)|\theta) \quad (1)$$

where $F_X(x)$ and $F_Y(y)$ are the distribution functions of random variables X and Y.

In this study, frequently used Archimedean copulas are considered to set up copula-based models. Let $C(.,.|\theta)$ be a bivariate parametric copula function and $(u,v) \in [0,1]x[0,1]$. Archimedean copulas are defined as $C(u,v) = \phi^{-1}(t)(\phi(u)+\phi(v))$ with generator function $\phi(t)$. Generator function $\phi : [0,1] \to [0,\infty]$ is a continuous, strictly decreasing and convex function with $\phi(1) = 0$ and $\phi^{-1}(t)$ shows the so called pseudo-inverse of generator function [15].

In some cases, Kendall's τ can be more applicative instead of the copula parameter θ in the sense of monotone transforms of copulas [8]. The closed forms of copula functions, the generator functions and the relationships between Kendall's τ and copula parameter θ for Archimedean copula family are given as follows [8], [15].

2.1. Clayton Copula

$$C(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}}, \qquad \theta \in [1,\infty)/\{0\} \quad (2)$$

$$\phi(t) = \frac{t^{-\theta} - 1}{\theta} \quad (3)$$

$$\tau = \frac{\theta}{\theta + 2} \quad (4)$$

2.2. Gumbel copula

$$C(u,v) = \exp\left[-\left((-\log u)^{\theta} + (-\log v)^{\theta}\right)^{\frac{1}{\theta}}\right], \qquad \theta \geq 1 \quad (5)$$

$$\phi(t) = (-\log t)^{\theta} \quad (6)$$

$$\tau = \frac{\theta - 1}{\theta} \quad (7)$$

2.3. Frank copula

$$C(u,v) = -\frac{1}{\theta}\log\left(1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{\left(e^{-\theta}-1\right)}\right), \quad \theta \in \Re/\{0\} \tag{8}$$

$$\phi(t) = -\ln\frac{-e^{-\theta t}-1}{e^{-\theta}-1} \tag{9}$$

$$\tau = 1+\frac{4}{\theta}\left(D_1\left(u,v|\theta\right)-1\right) \tag{10}$$

$D_1\left(u,v|\theta\right) = \dfrac{e^{\theta}\left(e^{\theta v}-1\right)}{e^{\theta(u+1)}+e^{\theta(v+1)}-e^{\theta}-e^{\theta(u+v)}}$ demostrates the first partial derivative.

3. The relationship between copula and aggregate loss

The aggregate loss is basically defined as the sum of monetary losses of all the claims. Aggregate loss can be calculated considering or ignoring the dependence between claim components.

3.1. Aggregate loss under the independence assumption

In risk theory, there are two main models: individual risk model and collective risk model. In the individual risk model, the claim frequency is assumed fixed. Otherwise, the aggregate loss has a compound distribution with the primary distribution being the claim frequency and the second one being the claim severity in collective risk model [16]. According to the collective risk model, claim severity and claim frequency are assumed independent random variables. Let Y displays the claim frequency and X_i shows the claim severity of the ith loss for i=1,..,Y. Aggregate loss which is denoted by S is defined as follows:

$$S = X_1 + \ldots + X_Y \tag{11}$$

X_i 's are assumed independently and identically distributed positive integer-valued random variables. Using the properties of compound distribution, the expected value and the variance of aggregate loss can be calculated as $E(S) = E(X)E(Y)$ and $Var(S) = E(Y)Var(X) + Var(Y)\left[E(X)\right]^2$, respectively.

In most studies expected value and the variance of aggregate loss is calculated without considering the distribution of aggregate loss due to the difficulty of obtaining closed form of the distribution function. The distribution of S can be obtained using convolution method, Panjer and De Peril recursions [16], [17]. Krämer et al. [8] proposed a new approach to define the distributions of policy loss and the aggregate loss considering dependency between claim severity and frequency using copula functions.

3.2. Estimation of aggregate loss using copula under the assumption of dependency

Aggregate loss can be estimated by copulas in the presence of dependency between claim components. The relationship between copula and aggregate loss can be explained by copula-based model relied on the distributions of claim components and the copula functions.

Let assume that claim severity and frequency have Gamma and zero-truncated Poisson distributions, respectively. Density and probability functions are given by Equation (12) and Equation (13) using mean parameterization. First moment, second moment and variance of distributions are also given by Table 1 as follows.

$$f_X\left(x \middle| \mu,\upsilon^2\right)=\frac{1}{\Gamma\left(\frac{1}{\upsilon^2}\right)}\left(\frac{1}{\mu\upsilon^2}\right)^{\frac{1}{\upsilon^2}} y^{\left(\frac{1}{\upsilon^2}\right)-1}\exp\left(-\frac{x}{\mu\upsilon^2}\right), \qquad x\geq 0 \tag{12}$$

$$f_Y\left(y \middle| \lambda\right)=\frac{\lambda^y e^{-\lambda}}{y!\left(1-e^{-\lambda}\right)}, \qquad y=1,2,3,... \tag{13}$$

Table 1. Moments of the Distributions of Claim Severity and Frequency

Variables	Distribution	First Moment	Second Moment	Variance
Claim Severity (X)	Gamma(μ,υ^2)	$E(X)=\mu$	$E(X^2)=\mu^2\left(1+\upsilon^2\right)$	$Var(X)=\mu^2\upsilon^2$
Claim Frequency (Y)	Zero-Truncated Poisson (λ)	$E(Y)=\frac{\lambda}{1-e^{-\lambda}}$	$E(Y^2)=\frac{\lambda+\lambda^2}{1-e^{-\lambda}}$	$Var(Y)=\frac{\lambda+\lambda^2}{1-e^{-\lambda}}-\left(\frac{\lambda}{1-e^{-\lambda}}\right)^2$

Let $C\left(.,.\middle|\theta\right)$ be a bivariate Archimedean copula where θ is the copula parameter. Joint cumulative distribution function and joint probability density function can be obtained Sklar Theorem [14] and by the mixed copula approach [3] as follows.

3.2.1. Joint cumulative distribution function of claim severity and frequency

The joint cumulative distribution function of claim severity and frequency can be written by using Sklar's Theorem as follows.

$$F_{XY}\left(x,y\middle|\mu,\upsilon^2,\lambda,\theta\right)=C\left(F_X(x),F_Y(y)\middle|\theta\right) \tag{14}$$

3.2.2. Joint probability density function of claim severity and frequency

The joint probability density function of claim severity and frequency can be written by the mixed copula approach which is proposed by Song [3]. The mixed copula approach, allows the usage of copula functions which are used with only the continuous random variables, also together with discrete random variables [7]. This approach uses the Radon-Nikodym Theorem to convert the discrete measurement to continuous measurement.

$$f_{XY}\left(x,y\middle|\mu,\upsilon^2,\lambda,\theta\right)=\frac{\partial}{\partial x}P\left(X\leq x,Y=y\right)=\frac{\partial}{\partial x}P\left(X\leq x,Y\leq y\right)-\frac{\partial}{\partial x}P\left(X\leq x,Y\leq y-1\right)$$

$$= \frac{\partial}{\partial x} C\left(F_X(x), F_Y(y) \middle| \theta\right) - \frac{\partial}{\partial x} C\left(F_X(x), F_Y(y-1) \middle| \theta\right)$$

$$f_{XY}\left(x, y \middle| \mu, \upsilon^2, \lambda, \theta\right) = f_X(x)\left[D_1\left(F_X(x), F_Y(y) \middle| \theta\right) - D_1\left(F_X(x), F_Y(y-1) \middle| \theta\right)\right] \tag{15}$$

$D_1(.,.|\theta)$ shows the first partial derivative and the Equation (15) can be proved by mixed copula approach. More detail information can be found in [3].

To indicate policy loss considering the dependence between the claim severity and frequency, Krämer et al. [8] define a new variable as L_i where X_i and Y_i display the ith claim severity and frequency, respectively.

$$L_i := X_i.Y_i \tag{16}$$

The expected value and the variance of the ith policy loss are denoted by μ_{L_i} and $\sigma^2_{L_i}$, respectively. Joint density function of loss can be written by mixed copula approach as follows [8].

$$f_L\left(l \middle| \mu, \upsilon^2, \lambda, \theta\right) = \sum_{y=1}^{n} \frac{1}{y} f_X\left(\frac{l}{y}\right)\left[D_1\left(F_X\left(\frac{l}{y}\right), F_Y(y) \middle| \theta\right) - D_1\left(F_X\left(\frac{l}{y}\right), F_Y(y-1) \middle| \theta\right)\right], \; l > 0 \tag{17}$$

Where there are n policies, aggregate loss S can be written as follows:

$$S := L_1 + \ldots + L_n = X_1.Y_1 + \ldots + X_n.Y_n \tag{18}$$

Since the policy loss L_i has positive and continuous distribution, aggregate loss has also positive and continuous distribution. Furthermore, according to the central limit theorem,

$$\frac{S - \sum_{i=1}^{n} \mu_{L_i}}{\frac{\sum_{i=1}^{n} \sigma^2_{L_i}}{\sqrt{n}}} \tag{19}$$

the asymptotic distribution of aggregate loss S is the standard normal distribution and can be displayed by $S \xrightarrow{D} N(0,1)$ using the Equation (19) [8].

4. Application

Analyses in the application part are carried out using the R packages “CopulaRegression” [18], “MASS” [19] and “VineCopula” [20]. “CopulaRegression” package works dependent on “MASS” and “VineCopula” packages. “CopulaRegression” package is established on the assumption of Gamma distributed claim severity and zero-truncated Poisson distributed claim frequency. A system is designed contained insured with Gamma-distributed claim severity and zero-truncated Poisson distributed claim frequency. For that purpose, first of all an one-year real comprehensive insurance data set taken from a non-life insurance company for year 2017 consisted of 2820 observations is analyzed to determine the parameters of distributions. Before the parameter estimation, the distributions of claim severity and the claim frequency are analyzed by Kolmogorov Simirnov Test using SPSS Statistics 23. As, Gamma distribution is formed by Exponential distribution and zero-truncated Poisson distribution is a special form of Poisson distribution, goodness of fit for severity and frequency is carried out for Exponential and Poisson. According to results of goodness of fit, claim severity and frequency of the real comprehensive insurance data fit aforementioned distributions. It is reasonable to using this data in the parameter estimation. To obtain the information of the claim severity, policies with at least one claim frequency are taken over. Parameters are estimated using the moments method. Details of the mean parametrization have been given by Table 1. Descriptive statistics are given by Table 2 for parameter estimation.

Table 2. Descriptive Statistics of Comprehensive Insurance Data

Variables	Minimum	Maximum	Mean	Median	Variance
Claim Severity (X)	51.24	35477.00	1759.4936	2933.58735	8605934.753
Claim Frequency (Y)	1	4	1.7034	0.4678	0.2190

Using moments method, mean parameter and dispersion parameter of Gamma distribution are approximately obtained as 1760 and 1, respectively and for zero-truncated Poisson distribution mean parameter is assumed as 1.18.

“CopulaRegression” package firstly simulate n observations claim severity from the marginal Gamma distribution, and for each severity, the package sample an observation from the conditional distribution of claim frequency given claim severity. The conditional density function for a Gamma distributed claim severity (X) and zero-truncated claim frequency (Y) can be written as $P(Y = y \mid X = x) = \dfrac{f_{XY}(x, y)}{f_X(x)}$ .In the second step, the package obtain the conditional distributions [18].

In this study, conditional probability mass function of the claim frequency is obtained as a first step to write copula-based model combining the marginal distributions and the copula function. While the effects of the copula on the conditional probability mass function of the claim frequency are demonstrated by Figure 1, the impressions of copula parameters on the conditional distribution are displayed by Figure 2. To investigate the effects of copula type, the conditional probability mass function of the claim frequency is plotted using Clayton, Gumbel and Frank copula functions with a fixed θ parameter. It is assumed that condition on claim severity of 2000 TL.

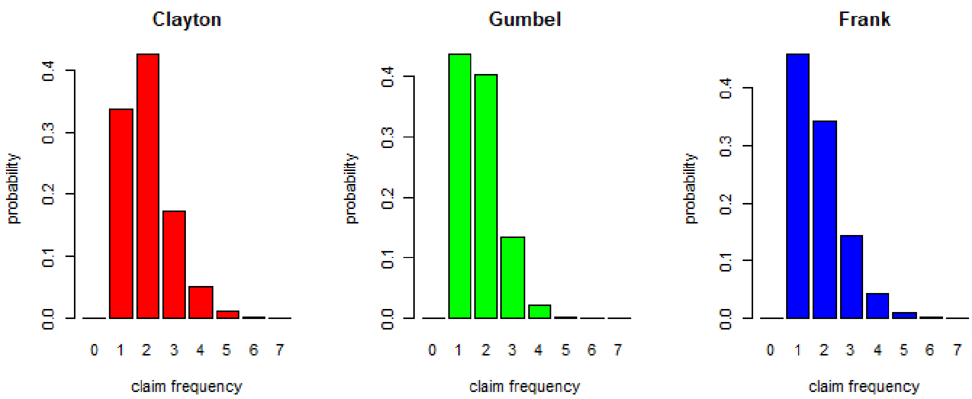

Figure 1. Conditional Probability Mass Function of the Claim Frequency (Clayton, Gumbel and Frank Copula Functions with a Fixed Parameter (θ=1.5))

According to the Figure 1, it is explicit that copula type influences the probability of claim frequency. The probabilities are higher for small values using Gumbel copula compared with other types of copula. The probability of one claim is higher with Gumbel and Frank copulas than the probability with Clayton. To investigate the effects of copula parameter, different parameters have been tried for all copula types. Instead of giving all results, the conditional probability mass functions of the claim frequency are plotted using different copula parameters for Clayton copula where parameter effect can be observed better.

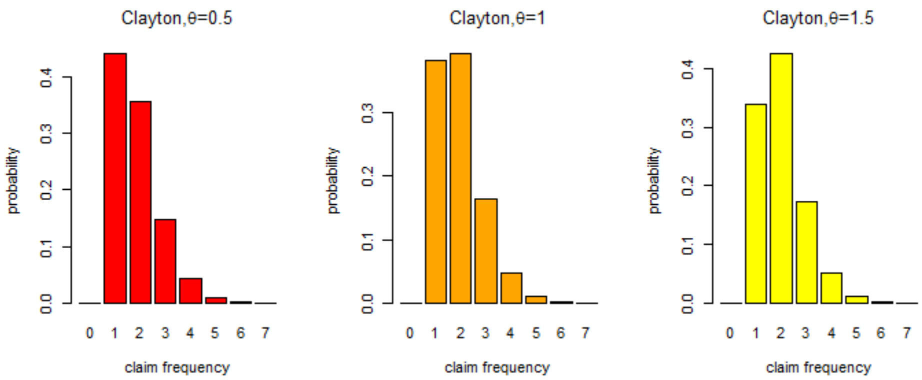

Figure 2. Conditional Probability Mass Function of the Claim Frequency (Clayton Copula Function with Different Parameters (θ=0.5; 1; 1.5))

It is observed from Figure 2, the higher values of θ, the greater conditional probability value of claim frequency, especially for smaller claims. It is deduced from Figure 1 and 2, dependence between claim frequency and severity can be modeled by Archimedean copulas flexibly.

Cumulative distribution function of claim severity and frequency given by Equation (14) are obtained with Clayton, Gumbel and Frank copulas for different parameters. Cumulative distribution functions using different copulas are plotted by Figure 3. It is noticed that the type of copula has no significant impact on the cumulative distribution function for small τ values. However, for larger τ values, the graph of cumulative distribution function using Frank copula is different the graphs using the other two copula types. Smaller cumulative probabilities are obtained for larger τ values while using Frank copula. Increasing curves are obtained in accordance with the general characteristics of the distribution functions.

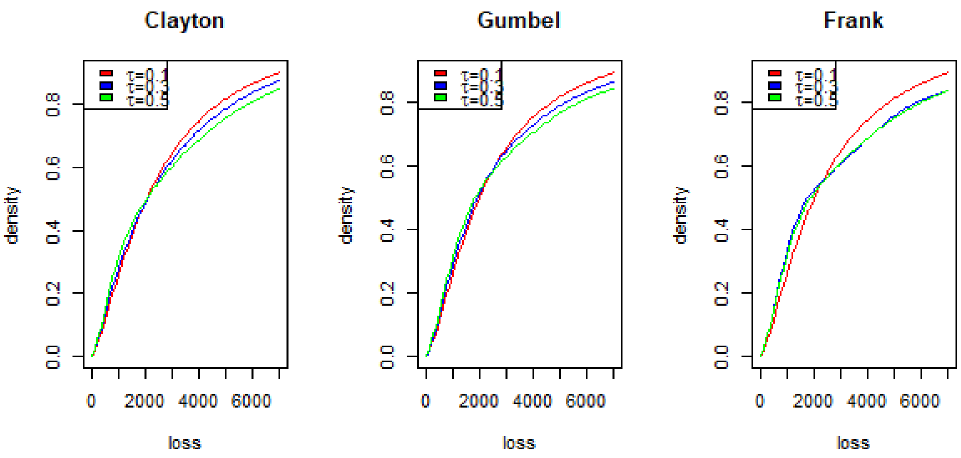

Figure 3. Cumulative Distribution Functions of Aggregate Loss for Different Copula Functions with Different Parameters

Joint probability density function of claim severity and frequency given by Equation (15) are obtained with Clayton, Gumbel and Frank copulas for different parameters. Joint density functions using different copulas are plotted by Figure 4.

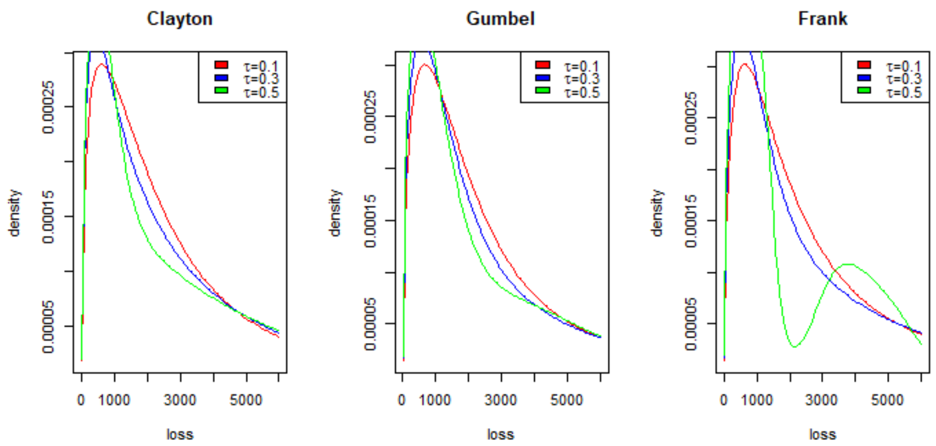

Figure 4. Probabilities of the Aggregate Loss for Different Copula Functions with Different Parameters

As the marginal Gamma distributed claim severity has a right-skewed distribution, the aggregate loss is also inclined to be right-skewed according to the Figure 4. While the probabilities are similar for Clayton and Gumbel copulas, Frank copula is more sensitive to the large value of the copula parameter. The results of the probability density function for Frank copula also support the results found for cumulative distribution function above.

Expected aggregate loss in the presence of dependence between claim components are estimated with the copula-based model. The estimated loss using different copulas and parameters are summarized with Table 3 as follows.

Table 3. Expected Loss using Clayton, Gumbel and Frank Copulas for Different Parameters

Type of Copula	τ	θ	Expected Loss
Clayton	0.1	0.2222222	2285.120
	0.2	0.5000000	2336.037
	0.3	0.8571429	2389.853
	0.4	1.3333333	2447.339
	0.5	2.0000000	2509.511
	0.6	3.0000000	2577.670
Gumbel	0.1	1.1111110	2372.550
	0.2	1.2500000	2492.395
	0.3	1.4285710	2596.252
	0.4	1.6666670	2684.361
	0.5	2.0000000	2757.008
	0.6	2.5000000	2814.559
Frank	0.1	0.9073682	2309.685
	0.2	1.8608840	2384.995
	0.3	2.9174345	2461.079
	0.4	4.1610643	2536.578
	0.5	5.7362827	2610.203
	0.6	7.9296423	2680.823

For Clayton, Gumbel and Frank copulas, same values of Kendall's τ parameter are taken and the corresponding values of θ copula parameter are given in Table 3. The higher values of τ and θ, the higher expected loss for all copulas. Expected losses are higher using Gumbel copulas compared the other copulas. Furthermore, some statistical values such as quantiles of expected loss are calculated and given by Table 4.

Table 4. Quantiles of Aggregate Loss using Clayton, Gumbel and Frank Copulas for τ =0.5

Type of Copula	%	Value at the quantile is evaluated	Expected Loss(τ=0.5)
Clayton	25	743.7378	2509.511
	50	1534.867	
	75	3211.736	
Gumbel	25	742.5345	2757.008
	50	1462.432	
	75	2917.698	
Frank	25	740.4356	2610.203
	50	1472.067	
	75	3156.319	

Quantiles of aggregate loss given by Table 4 can be useful for interval estimation of aggregate loss under dependence assumption. The values of expected aggregate loss are ranked between the 50 % and 75 % quartiles. It confirms that the distribution of aggregate loss under dependency assumption is also right skewed.

5. Concluding Remarks

In non-life insurance mathematics claim severity and frequency are generally assumed independent and the aggregate loss is over or under estimated compared the actual aggregate loss under this assumption. Dependency between claim components can taken into consideration using copula-based model approach proposed by Kramer et al. [8]. Copula-based models can be created by bivariate Archimedean copula family and mixed vairables such as continuous claim severity and discrete claim frequency. Mixed variables can be included in the same copula function with the mixed copula approach. In this study, joint cumulative distribution and joint probability density functions of claim severity and frequecy are obtained using the copula-based model. The aggregate loss is modeled using different bivariate Archimedean copula functions and different copula parameters. The effects of the parameters and the type of copulas on dependence modeling are analysed and interprered.

It has been noticed that copula-based models allow for flexible dependency modeling of claim severity and frequency in non-life insurance mathematics. Flexibility arises from the use of different copula functions with different parameters in these models. According to the results of the study, Frank copula is more sensitive to dependence modeling in cases where the relationship between the claim severity and frequency is high. It has been observed that aggregate loss increases as τ value increases for all Archimedean copula functions used. However, higher aggregate loss is estimated when the dependency is modeled with Gumbel copula. Furthermore, the quantiles of aggregate loss that can be used in interval estimation of aggregate loss under dependence assumption are calculated with the help of the R packages "CopulaRegression". It has been seen that the values of expected aggregate loss are ranked between the 50 % and 75 % quartiles, which is an expected result since the distribution of aggregate loss under dependency assumption is also right skewed.

Copula-based models in this study can be developed as copula-based regression models including generalized linear models and the results of models can be compared. This study can be enhanced by using other parametric copulas such as Gauss and t-copulas or non-parametric copulas such as Bernstein copulas.

References

[1] F. Lundberg, 1903, *I. Approximerad framstallning af sannolikhetsfunktionen: II. Aterforsakring af kollektivrisker*. Almqvist & Wiksell.

[2] E. W. Frees, E. A. Valdez, 1998, Understanding relationships using copulas. *North American actuarial journal*, *2*(1), 1-25

[3] X. K. Song, P. X. K. Song, 2007, *Correlated data analysis: modeling, analytics, and applications*. Springer Science & Business Media.

[4] R. Kastenmeier, 2008, Joint regression analysis of insurance claims and claim sizes, Technische Universitat München, Mathematical Sciences, Diploma Thesis.

[5] P. X. K. Song, M. Li, Y. Yuan, 2009, Joint regression analysis of correlated data using Gaussian copulas, *Biometrics*, *65*(1), 60-68.

[6] N. Kolev, D. Paiva, 2009, Copula-based regression models: A survey. *Journal of statistical planning and inference*, *139*(11), 3847-3856.

[7] C. Czado, R. Kastenmeier, E. C. Brechmann, A. Min, 2012, A mixed copula model for insurance claims and claim sizes, *Scandinavian Actuarial Journal*, 4, 278-305.

[8] N. Kramer, E. C. Brechmann, D. Silvestrini, C. Czado, 2013, Total loss estimation using copula-based regression models. *Insurance: Mathematics and Economics*, *53*(3), 829-839.

[9] S. Gschlößl, C. Czado, 2007, Spatial modelling of claim frequency and claim size in non-life insurance. *Scandinavian Actuarial Journal*, 3, 202-225.

[10] J. Garrido, C. Genest, J. Schulz, 2016, Generalized linear models for dependent frequency and severity of insurance claims. *Insurance: Mathematics and Economics*, *70*, 205-215.

[11] H. Arvidson, S. Francke, 2007, Dependence in non-life insurance, UUDM Project Report, 23.

[12] S. Anastasiadis, S. Chukova, 2012, Multivariate insurance models: an overview. *Insurance: Mathematics and Economics*, *51*(1), 222-227.

[13] S. Eryilmaz, S., 2017, On compound sums under dependence. *Insurance: Mathematics and Economics*, *72*, 228-234.

[14] A. Sklar, 1959, Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de L'Université de Paris*, Vol. 8, pp. 229-231.

[15] R. B. Nelsen, 2006, *An introduction to copulas*. Springer Science & Business Media.

[16] Y. K. Tse, 2009, *Nonlife Actuarial Models: Theory, Methods and Evaluation*, Cambridge University Press.

[17] S. A. Klugman, H. H. Panjer, G. E. Willmot, 2008, *Loss models: from data to decisions*, John Wiley & Sons.

[18] N. Kramer, D. Silvestrini, M. N. Kraemer, 2013, Package 'CopulaRegression'.

[19] B. Ripley, B. Venables, D. M. Bates, K. Hornik, A. Gebhardt, D. Firth, M.B. Ripley, 2013, Package 'mass'. Cran R.

[20] U. Schepsmeier, J. Stoeber, E. C. Brechmann, B. Graeler, T. Nagler, T. Erhardt, 2012, VineCopula: Statistical inference of vine copulas. R package version 1.