Assessing High-Risk Scenarios by Full-Range Tail Dependence Copulas

Sponsored by
CAS, CIA, SOA Joint Risk Management Section

Prepared By
Lei Hua
Michelle Xia
December 2013

Assessing high-risk scenarios by full-range tail dependence copulas*

Lei Hua† Michelle Xia‡

December 12, 2013

Abstract. Copulas with a full-range tail dependence property can cover the widest range of positive dependence in the tail, so that a regression model can be built accounting for dynamic tail dependence patterns between variables. We propose a model that incorporates both regression on each marginal of bivariate response variables and regression on the dependence parameter for the response variables. ACIG copula that possesses the full-range tail dependence property is implemented in the regression analysis. Comparisons between regression analysis based on ACIG and Gumbel copulas are conducted, showing that ACIG is generally better than Gumbel copula when there is intermediate upper tail dependence. A simulation study is conducted to illustrate that dynamic tail dependence structures between loss and ALAE can be captured by using the one-parameter ACIG copula. Finally, we apply the ACIG and Gumbel regression models respectively for a dataset from the Medical Expenditure Panel Survey of the United States. The empirical analysis suggests that the regression model with the ACIG copula improves the assessment of high-risk scenarios, especially for aggregated dependent risks.

Key words: ACIG copula, tail order, dynamic dependence, ALAE, MEPS dataset, regression analysis.

*The research conducted was sponsored by the Society of Actuaries (SOA), Casualty Actuarial Society (CAS) and Canadian Institute of Actuaries (CIA). The financial support from SOA, CAS and CIA, and the precious time of the volunteers involved in proposal reviews and project oversights is greatly appreciated. The first author is thankful to Dr. Peng Shi for some helpful discussion about the project. All remaining errors are our own.

†Corresponding author, hua@math.niu.edu, Division of Statistics, Northern Illinois University, DeKalb, IL, 60115, United States.

‡cxia@niu.edu, Division of Statistics, Northern Illinois University, DeKalb, IL, 60115, United States.

1 Introduction

Regression analysis with a variety of generalized linear models (GLMs, which include generalized additive models in our following discussion) has been proven useful in making statistical inference on univariate distributions. In particular, it has been widely used in property and casualty insurance companies for rate making. However, those models are mainly for pricing, and may not be suitable for quantitative risk management. The main reason is that the pricing models are chosen and the parameters are calibrated based on the explanatory variables that take values having relatively higher possibilities. However, extreme values, although having relatively smaller chances to occur, may have a direct impact on financial stability. In the field of quantitative risk management, the study on tail behavior of loss distributions has attracted more attention than ever since the most recent financial crisis starting from 2007. Among many debates about the financial crisis, limitations of Gaussian copulas have also been discussed. We refer to Donnelly and Embrechts (2010) for the drawbacks of using Gaussian copula for modeling credit risks.

All statistical models have certain limitations. For quantitative risk management, one major limitation of the Gaussian copula is that it does not well capture the joint tail behavior of large risks when there actually are clusters of large losses. More precisely, the tail dependence parameter $\lambda = \lim_{x\to\infty} \mathbb{P}[X_1 > x | X_2 > x] = 0$, where X_1 and X_2 are identically distributed random variables and have the Gaussian copula as their dependence structure. The lack of tail dependence can be easily fixed by using a copula that has upper tail dependence such as the Gumbel copula and the student t copula, of which the upper tail dependence parameter $\lambda > 0$. However, one may overreact to the issue by using copulas that have tail dependence when the actual dependence in the upper tail is not strong enough. Therefore, we need a copula that can capture both cases where the upper tail dependence parameter can be either $\lambda > 0$ or $\lambda = 0$.

When a copula family is able to account for both $\lambda > 0$ and $\lambda = 0$ for positive dependence in the tail, we refer to it as a full-range tail dependence copula; note that, here the "tail dependence" is a generic description of dependence in the tail, not the concept discussed in Section 2.1.10 of Joe (1997). For the latter, we will use "usual tail dependence" specifically for the case where $\lambda > 0$. When $\lambda > 0$, it is the usual tail dependence case, and the value of λ itself can be used to quantify the degree of dependence. However, when $\lambda = 0$, we need more information to quantify the degree of dependence. In Hua and Joe (2011), a notion of "tail order" is used to quantify the degree of dependence in the tails for the case where $\lambda = 0$. Then a concept of "intermediate tail dependence" has been proposed to account for the positive tail dependence with $\lambda = 0$.

The importance of using a full-range tail dependence copula in assessing high-risk scenarios is also suggested by the findings in Hua and Joe (2014). In a regression setting, if one wants to conduct inference on the risk measures of the following forms $\mathbb{E}(Y_1|Y_2 > t, \mathbf{X} = \mathbf{x})$ or $\mathbb{E}(Y_1|Y_2 = t, \mathbf{X} = \mathbf{x})$, where t is Value at Risk (VaR) for Y_2, then the strength of tail dependence between Y_1 and Y_2 conditioning on $\mathbf{X} = \mathbf{x}$ becomes very important, because the risk measures are very sensitive to the strength of tail dependence. Under such a situation, a full-range tail dependence copula may significantly improve the modeling.

The goal of the paper is to study how a full-range tail dependence copula would be useful in assessing tail risks in a regression setting. We will first introduce the concepts of tail order and full-range tail dependence copulas, and a one-parameter full-range tail dependence copula that is referred to as ACIG copula will be studied. In particular, we will compare the ACIG copula with the commonly-used Gumbel copula, and our emphasis will be the comparison of different upper tail behavior of these two copulas. We will illustrate that the ACIG copula has a wider range of tail dependence than the Gumbel copula, and the ACIG copula has less cross-validated prediction errors than the Gumbel copula under a simulation study. Then, we will develop a copula-based regression model where the response variable is bivariate and the dependence degrees for the bivariate response variable is allowed to change according to the values of explanatory variables. A simulated loss-and-expense dataset will be used to demonstrate the dynamic dependence structures between response variables and how full-range tail dependence copula can be used in capturing the changing patterns of tail dependence. Finally, we will conduct an empirical analysis for a Medical Expenditure Panel Survey dataset, for which the full-range tail dependence ACIG copula can improve the modeling for assessing high-risk scenarios.

The main contribution of the paper is the following: the concept of full-range tail dependence is proposed for modeling dynamic tail dependence patterns, which overcomes the problem of the existing models being either only tail independent or only tail dependent, and a one-parameter full-range tail dependence copula has been studied and implemented in regression models that can improve the modeling for assessing high-risk scenarios. The regression model with a full-range tail dependence copula can improve risk assessments for not only marginal but also aggregated losses.

The paper is organized as the following: In Section 2, basic concepts of tail order and full-range tail dependence copula will be introduced. In Section 2.3, we use a cross-validated prediction error to compare the model performance between ACIG and Gumbel copulas. Regression models using a full-range tail dependence copula will be constructed in Section 3. A simulation study using the ACIG copula for modeling the dynamic dependence structures between auto insurance losses and expenses is reported in Section 4. In Section 5, we use a Medical Expenditure Panel Survey dataset to demonstrate the implemented regression model with the ACIG copula, and its comparison to the Gumbel model. Section 6 will conclude the paper, and some technical arguments for implementing the models are reported in Section 7.

2 Tail order and full-range tail dependence

2.1 Full-range tail dependence

For a random vector $(X_1, \ldots, X_d)$ that has a joint cumulative distribution function (cdf) F and univariate cdf $F_i, i = 1, \ldots, d$, there is a copula function $C : [0,1]^d \to [0,1]$ such that for any $\mathbf{x} = (x_1, \ldots, x_d)$ on the support of F, $F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$. When F_i's are continuous, the copula C is uniquely determined. Use $\widehat{C}$ as the corresponding survival copula of C; that is, $\widehat{C}(u_1, \ldots, u_d) = \overline{C}(1 - u_1, \ldots, 1 - u_d)$, where $\overline{C}$ is the survival function of C, and defined as $\overline{C}(u_1, \ldots, u_d) := 1 + \sum_{\emptyset \neq I \subseteq I_d} (-1)^{|I|} C_I(u_i, i \in I)$ with C_I the copula for the I-marginal, $I_d := \{1, \ldots, d\}$, and $|I|$ the number of elements in the set I; when $|I| = 1$, the notation $C_I(u_i) = u_i, i \in I$.

The upper tail behavior of C is simply the lower tail behavior of $\widehat{C}$, and vice verse. Since the upper tail is usually relevant for risk assessment if the random variables represent amounts of losses, we are only interested in the upper tail of a copula in this paper. Therefore, in what follows, we consider the lower tail of $\widehat{C}$ when defining relevant concepts about the tail, and similar concepts can also be defined for the upper tail of $\widehat{C}$.

If $\widehat{C}(u, \ldots, u) \sim \lambda u$ with $0 < \lambda \leq 1$, as $u \to 0^+$, then C has usual upper tail dependence with the dependence parameter being λ, where "$\sim$" means asymptotically equivalence; that is, $g(x) \sim h(x), x \to x_0 \iff \lim_{x \to x_0} g(x)/h(x) = 1$. If $\widehat{C}(u, \ldots, u) \sim u^\kappa \ell(u)$, as $u \to 0^+$, with ℓ a slowly varying function[1] and $1 \leq \kappa$, then κ is referred to as the upper tail order of the copula C. Clearly, if $\widehat{C}$ has usual upper tail dependence, then its upper tail order $\kappa = 1$ and $\lambda = \lim_{u \to 0^+} \ell(u)$. A larger value of κ indicates a weaker dependence in the tail. When there is usual tail dependence, that is, when $\lambda > 0$, we use the value of λ to quantify the degree of dependence in the tail; when $\kappa > 1$, thus $\lambda = 0$, we use the value of κ to quantify the degree of dependence in the tail. When $1 < \kappa < d$, the corresponding tail of C is said to have intermediate tail dependence with certain regularity conditions. For details about the notion of tail order, we refer the interested reader to Hua and Joe (2011, 2013).

When a permutation symmetric copula family C is able to account for $1 \leq \kappa \leq d$, we refer to it as a full-range tail dependence copula. For example, for a bivariate Gaussian copula with correlation coefficient $|\varrho| \neq 1$, the tail order $1 < \kappa = 2/(1+\varrho) < \infty$ (Hua and Joe, 2011, Example 1), and it is not a full-range tail dependence copula because $\kappa \neq 1$. In the existing

[1] A measurable function g is said to be slowly varying at x_0 if for any $t > 0$, $\lim_{x \to x_0} g(tx)/g(x) = 1$. For example, $\log(x)$ is slowly varying at the infinity.

parametric copula families, there are only a few non-trivial copula families that have the full-range tail dependence property. One example is the Archimedean copula constructed by the Laplace Transform (LT) of an inverse Gamma distribution (referred to as ACIG copula, (Hua and Joe, 2011, Example 4)). Another example is the Archimedean copula constructed by the mixture of generalized Gamma and simplex (referred to as GGS copula, (Hua, 2013, Example 2)).

In the next two subsections, we will focus on the ACIG copula and its comparisons with the Gumbel copula, because the ACIG will be more useful than the GGS copula for modeling the dependence structures for the datasets to be analyzed in this paper. We refer to Hua (2013) for details about the GGS copula and its applications in modeling the dependence between loss frequency and loss severity.

2.2 ACIG copula

An Archimedean copula can be constructed by an Archimedean generator ψ as the following.

$$C(u_1, \ldots, u_d) = \psi(\psi^{-1}(u_d) + \cdots + \psi^{-1}(u_d)), \quad u_i \in [0,1], \quad i = 1, \ldots, d.$$

The Archimedean generator ψ can often be chosen as an LT of a positive random variable; that is, $\psi(s) = \int_0^\infty e^{-sx} F(dx)$, where F is the cdf of a positive random variable X. Then, $\psi(0) = 1$, $\psi(\infty) = 0$, and ψ is completely monotone. We refer to Joe (1997) for details about constructing Archimedean copulas based on the LT of a positive random variable.

Based on Example 4 of Hua and Joe (2011), let $Y = X^{-1}$, and X be distributed as Gamma$(\alpha, 1)$, where α is the shape parameter. Then the LT of the inverse Gamma random variable Y is

$$\psi(s;\alpha) = \frac{2}{\Gamma(\alpha)} s^{\alpha/2} K_\alpha(2\sqrt{s}), \quad s \geq 0, \ \alpha > 0, \tag{1}$$

where K_α is the modified Bessel function of the second kind. For the bivariate copula $C(u,v) := \psi(\psi^{-1}(u) + \psi^{-1}(v))$, the density is

$$c(u,v) = \psi''(\psi^{-1}(u) + \psi^{-1}(v)) \cdot [\psi'(\psi^{-1}(u))]^{-1} \cdot [\psi'(\psi^{-1}(v))]^{-1}. \tag{2}$$

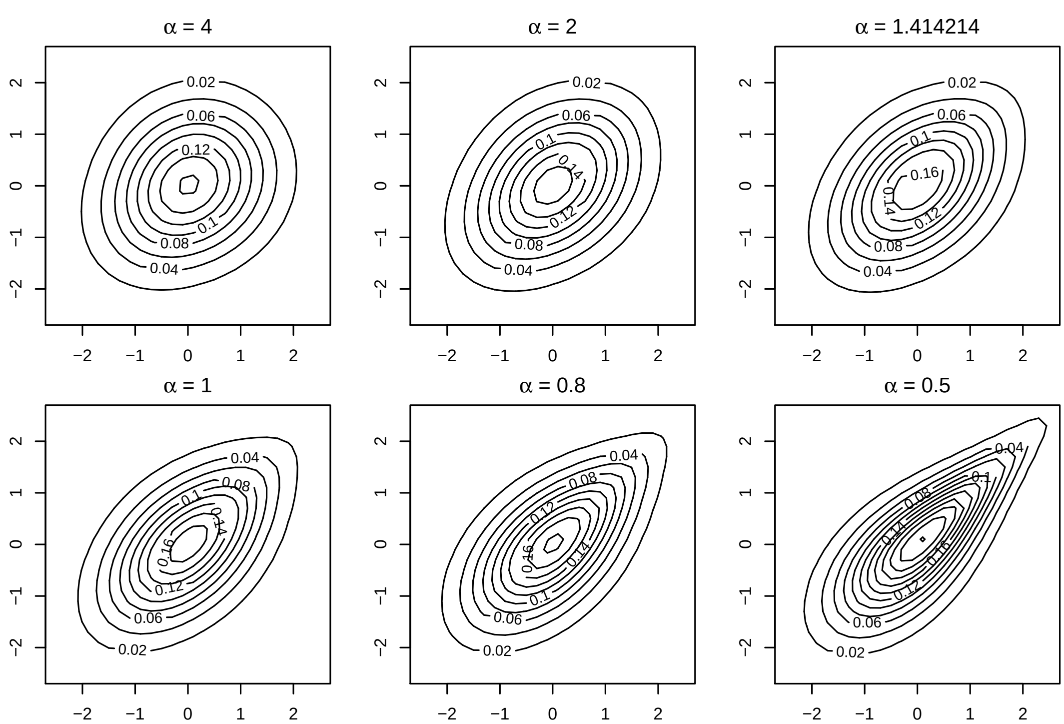
In order to calculate the density, we need to find ψ', ψ'' and ψ^{-1}. The first two can be obtained analytically as the following, and ψ^{-1} can be evaluated numerically.

$$\psi'(s) = -2s^{(\alpha-1)/2}K_{\alpha-1}(2\sqrt{s})/\Gamma(\alpha)$$
$$\psi''(s) = 2s^{(\alpha-2)/2}K_{\alpha-2}(2\sqrt{s})/\Gamma(\alpha)$$

Here we comment that, for data analysis with Archimedean copulas that are constructed by strictly decreasing ψ, as long as derivatives of the generator ψ can be obtained analytically, the copula is often useful for real applications, because a numerical method is usually very efficient and fast for getting ψ^{-1} for such a strictly decreasing function ψ.

Figure 1 illustrates the normalized contour plots of the ACIG copula. To get the normalized contour plots, we transform each copula marginal separately to the standard Normal; that is, transforming (u, v) by $(\Phi^{-1}(u), \Phi^{-1}(v))$, and then plotting the pairs of the latter, where Φ is the cdf of the standard Normal distribution. Based on Figure 1, we can clearly notice that ACIG copula is able to incorporate a much wider range of dependence in the upper tail with a single dependence parameter α. The nice property of this copula provides a suitable tool to modeling dynamic dependence between loss and expense, conditioning on a variety of values of the covariates.

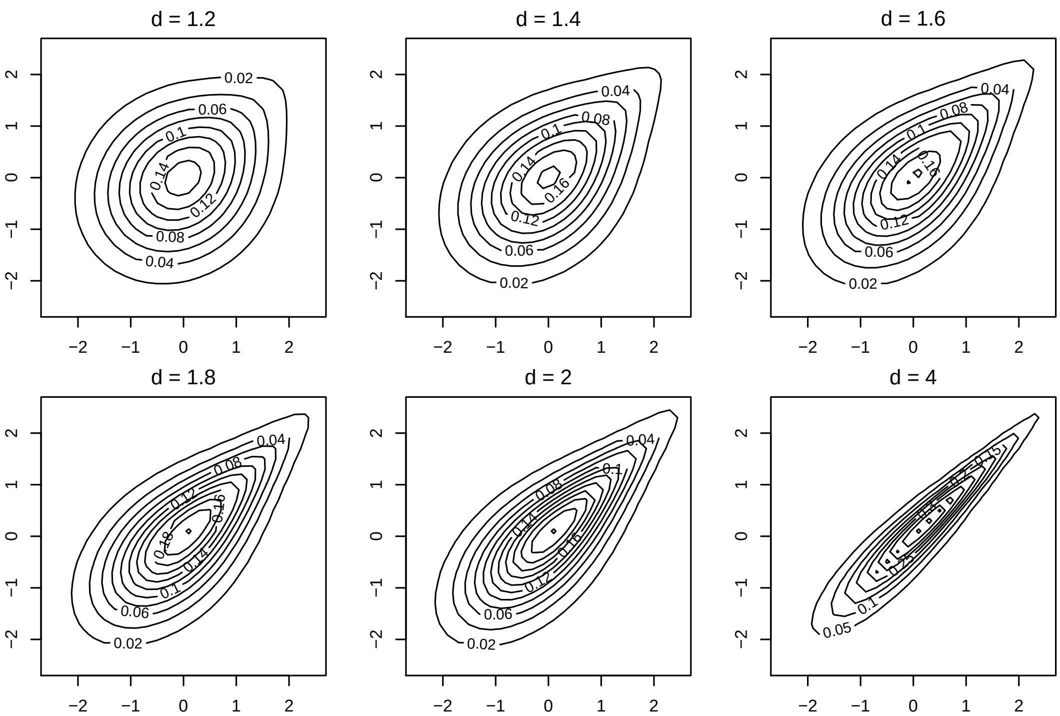
Figure 1: Normalized contour plots of ACIG copula



2.3 Comparisons

For the well known loss and ALAE dataset studied in papers such as Frees and Valdez (1998), and Klugman and Parsa (1999), Gumbel copula can be used to fit the dependence structures. Figure 2 shows normalized contour plots for Gumbel copula. Based on the contour plots for ACIG and Gumbel copulas, when there is usual upper tail dependence, the two copulas look very similar; however, when the dependence becomes weaker, the Gumbel copula still has a hump in the upper tail while the upper tail of ACIG copula is able to incorporate a relatively weaker dependence structure.

Figure 2: Normalized contour plots of Gumbel copula



In order to compare these two copulas regarding their capacity for fitting a variety of dependence in the upper tail, we apply the cross-validated prediction errors (CVPE) proposed in Acar et al. (2011) as a criteria. For a random sample $(u_i, v_i), i = 1, \dots, n$, where $0 \leq u_i, v_i \leq 1$, we use copula C to fit the sample. The leave-one-out CVPE for the copula C is then defined as follows.

$$\mathrm{CVPE}(C) = \sum_{i=1}^{n} \left[\left\{ u_i - \hat{\mathbb{E}}^{(-i)}(U_i|v_i) \right\}^2 + \left\{ v_i - \hat{\mathbb{E}}^{(-i)}(V_i|u_i) \right\}^2 \right],$$

where

$$\hat{\mathbb{E}}^{(-i)}(U_i|v_i) = \int_0^1 uc(u, v_i|\hat{\theta}^{(-i)})du, \quad i = 1, \ldots, n,$$

with $c(\cdot, \cdot|\hat{\theta}^{(-i)})$ the estimated copula density when the ith observation (u_i, v_i) is removed.

Now we are comparing the performance of these two copulas for fitting simulated datasets generated from ACIG and Gumbel copulas, respectively. Three different levels of dependence are considered in terms of Blomqvist's β; a higher value of β indicates a stronger positive dependence. Blomqvist's β of a bivariate copula C is defined as $\beta = 4 \times C(1/2, 1/2) - 1$. Then $\beta_1 = 0.1$, $\beta_2 = 0.3$, and $\beta_3 = 0.5$ are considered, and the corresponding values of parameters for the ACIG copula are $4.66639, 1.20143, 0.50171$, and for the Gumbel copulas are $1.11453, 1.43406, 1.99664$, respectively. We generate datasets from Gumbel copulas (denoted as $M_0 =$ G), and then fit the datasets with Gumbel and ACIG copulas (denoted as G and A) and calculate the leave-one-out CVPEs, respectively. We also generate datasets from ACIG copulas, and then fit the datasets with Gumbel and ACIG copulas and calculate the CVPEs, respectively again. We repeat the procedures 50 times, and the means and standard deviations of those CVPEs are reported in Table 1 for small sample size with $n = 100$, and for large sample size with $n = 1,000$, respectively.

Table 1: Cross-validated prediction errors (CVPE) for ACIG and Gumbel copulas

		$n = 100$				$n = 1000$			
		$M_0 =$ G		$M_0 =$ A		$M_0 =$ G		$M_0 =$ A	
Blomqvist β		G	A	G	A	G	A	G	A
$\beta_1 = 0.1$	CVPE	12.02	11.65	13.89	12.66	120.16	118.09	134.43	122.77
	s.d.	2.45	3.19	2.15	2.34	8.35	12.82	8.94	8.64
$\beta_2 = 0.3$	CVPE	5.55	4.66	6.28	5.48	55.56	45.70	65.06	56.03
	s.d.	1.29	1.07	1.51	1.33	4.63	3.72	4.83	4.30
$\beta_3 = 0.5$	CVPE	1.90	1.88	1.82	1.83	19.19	19.14	19.55	19.36
	s.d.	0.49	0.36	0.51	0.37	1.89	1.35	1.92	1.47

Based on Table 1, ACIG copula is generally better or comparable to Gumbel copula. For the ACIG copula, $\beta_1 = 0.1, 0.3, 0.5$ correspond to tail quadrant independence, intermediate tail dependence and usual tail dependence, respectively. When dependence in the upper tail is not as strong as the usual tail dependence case, ACIG copula generally outperforms Gumbel copula in the sense of the CVPE. In particular, for the intermediate tail dependence case, ACIG copula is clearly better. As expected, when the dependence in the upper tail is relatively stronger, it is hard to distinguish the ACIG copula from the Gumbel copula.

3 Regression for dependence

3.1 Dynamic dependence

Dependence modeling between loss and expense has been a popular research topic in actuarial science since the publication of Frees and Valdez (1998), and Klugman and Parsa (1999). Copulas have been applied to model such non-Gaussian dependence relationships between loss and certain associated expenses, although there were no covariates involved probably due to the lack of relevant datasets. For regression on bivariate response variables, the dependence parameter for the response variable is considered by papers such as de Leon and Wu (2011), and Czado et al. (2011); the latter models the dependence between loss frequency and loss severity, where the marginals are assumed to depend on the values of covariates but the dependence parameter is assumed to be homogeneous. Acar et al. (2011) applies a nonparametric approach for calibrating the dependence parameters according to the covariates, where the dependence parameter is allowed to change along the covariates.

Through a preliminary data analysis, we observe that if the dependence patterns conditioning on different values of covariates appear to be very different, then the model fitting could be improved by adding a regression analysis that accounts for the dynamic dependence patterns. From bodily injury loss (LOSS) and its allocated loss adjustment expense (ALAE) of a certain car insurance dataset, we noticed that the dependence structures between LOSS and ALAE change significantly for different duration of investigation of the loss events. When the duration of investigation is longer, the dependence in the upper tail becomes stronger. It makes sense as a longer duration of investigation may indicate a more complicated loss event and investigation, and thus more expense such as attorney fees could be involved, so that the dependence between LOSS and ALAE becomes stronger. Due to confidentiality, we only use a simulated dataset in Figure 3 to illustrate the idea, and the patterns look similar to real datasets. An empirical analysis based on real dataset will be conducted in Section 5, where the dataset is from the Medical Expenditure Panel Survey of the United States.

The changing dependence pattern indicates that the local dependence structure in the upper tail may cover a very wide range from independence to stronger positive dependence. However, the commonly used copula families such as the Gumbel copula can only account for a very narrow range of tail dependence pattern, that is the usual tail dependence. To overcome this obstacle, we need to consider a copula family that has a more flexible upper tail with less dependence parameters that can be linked to the covariates.

3.2 Regression models

The proposed model generally can be used to model response variables of any dimension. Now we use a bivariate case to illustrate the idea. Conditional marginal distributions of response variables Y_1 and Y_2 are modeled by GLMs or other univariate regression models, respectively. They are denoted as $F_1(\cdot|\mathbf{X}_1 = \mathbf{x}_1)$ and $F_2(\cdot|\mathbf{X}_2 = \mathbf{x}_2)$, respectively. Note that, the set of covariates $\mathbf{X}_1$ and $\mathbf{X}_2$ are not necessarily the same. But the covariates for the dependence parameter should also be contained in the covariates for both response variables. In what follows, we use the same set of covariates $\mathbf{X}$ for notational convenience.

Let $Y_i|\mathbf{X} = \mathbf{x}, i = 1, 2$ be continuous random variables conditioning on $\mathbf{X} = \mathbf{x}$ with cumulative distribution functions $F_{i|\mathbf{X}}(\cdot|\mathbf{x}), i = 1, 2$. Then the copula of those $Y_i|\mathbf{X} = \mathbf{x}, i = 1, 2$ is referred to as the conditional copula of $Y_i's$ given the covariates $\mathbf{X} = \mathbf{x}$. A bivariate one-parameter copula such as the ACIG copula that has full-range upper tail dependence will be used to model the conditional dependence between the response variables. Then the conditional joint cumulative distribution function is

$$F(y_1, y_2|\alpha(\mathbf{x}), \boldsymbol{\theta}_1(\mathbf{x}), \boldsymbol{\theta}_2(\mathbf{x})) = C(F_1(y_1|\boldsymbol{\theta}_1(\mathbf{x})), F_2(y_2|\boldsymbol{\theta}_2(\mathbf{x}))|\alpha(\mathbf{x})). \tag{3}$$

where $C(\cdot, \cdot|\alpha(\mathbf{x}))$ is the conditional copula, $\boldsymbol{\theta}_i(\mathbf{x}), i = 1, 2$ are the parameters for the two marginals respectively, and $\alpha(\mathbf{x})$ is the dependence parameter that is also linked to the covariates $\mathbf{x}$.

Since the dependence parameter α of ACIG copula is positive, we use the $\log(\cdot)$ as the link function. Moreover, certain nonlinear functionals such as natural cubic splines (ns) will be applied for continuous covariates as we do not know the functional relationship between the covariates and the dependence parameter α. That is,

$$\log(\alpha(\mathbf{x})) = \sum_{i=1}^{k} \mathrm{ns}_i(x_i) + \gamma_0 + \sum_{i=k+1}^{p} \gamma_i x_i,$$

where $x_1, \ldots, x_k$ are continuous covariates, $x_{k+1}, \ldots, x_p$ are categorical covariates. The forms of $\boldsymbol{\theta_i}(\mathbf{x}), i = 1, 2$ depend on how the marginal regression models are chosen.

With specified marginal parametric distributions F_1, F_2, and parametric copula C, we can get the density function of (3) as the following

$$f(y_1, y_2|\alpha(\mathbf{x}), \boldsymbol{\theta}_1(\mathbf{x}), \boldsymbol{\theta}_2(\mathbf{x})) = c(F_1(y_1|\boldsymbol{\theta}_1(\mathbf{x})), F_2(y_2|\boldsymbol{\theta}_2(\mathbf{x}))|\alpha(\mathbf{x})) \cdot f_1(y_1|\boldsymbol{\theta}_1(\mathbf{x})) \cdot f_2(y_2|\boldsymbol{\theta}_2(\mathbf{x})), \tag{4}$$

where c, f_1 and f_2 are the corresponding conditional density functions of the copula and the marginals. Then, the coefficients β's associated with the natural splines (ns) and the coefficients γ_i's for the discrete variables can be estimated by maximizing the following full likelihood function.

$$L(\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \alpha | y_1, y_2, \mathbf{x}) = \prod_{i=1}^{n} f(y_{1i}, y_{2i} | \alpha(\mathbf{x}_i), \boldsymbol{\theta_1}(\mathbf{x}_i), \boldsymbol{\theta_2}(\mathbf{x}_i)),$$

where $(y_{1i}, y_{2i}, \mathbf{x}_i)$ is the ith observation of the random sample of size n.

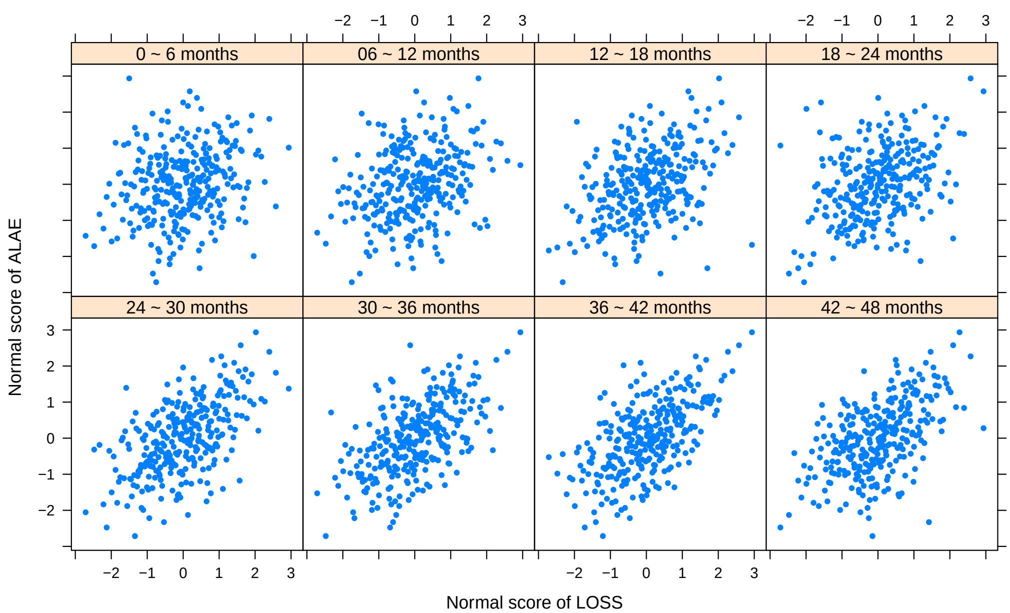
4 Simulation study

In this section, we apply the regression model mentioned in Section 3.2 for a simulated loss-and-ALAE dataset to illustrate the idea of using a full-range tail dependence copula to account for the dynamic dependence between LOSS and ALAE along covariates.

We generate a random sample of size $n = 2400$ from the ACIG copula, and the sample size for each month is 50. The dependence parameter is assumed to be a function of the claim duration in months ($x \in \{1, \ldots, 48\}$) according to the following formula, and Figure 3 is the normalized scatter plots for each period of 6 months.

$$\alpha(x) = 4.62 - 3.60 \times 10^{-1}x + 1.33 \times 10^{-2}x^2 - 2.30 \times 10^{-4}x^3 + 1.54 \times 10^{-6}x^4. \qquad (5)$$

Figure 3: Dependence in the upper tail becomes stronger as duration increases



We now link the dependence parameter α of the ACIG copula to the covariate "Duration" X. Since α has to be positive, we use log() as the link function. Natural cubic spline is used for the duration as we do not know the functional relationship between α and X. Then,

$$\log(\alpha(x)) = \text{ns}(x) = \beta_1 b_1(x) + \cdots + \beta_p b_p(x),$$

where $b_1(x), \ldots, b_p(x)$ are the spline basis and β_i's are the coefficients to be estimated. Let $u_i, v_i, i = 1, \ldots, n$ be the random sample generated by the ACIG copula. The likelihood function L and the loglikelihood function l in terms of the Archimedean generator ψ are the following.

$$L(\beta_1, \ldots, \beta_p | u_i, v_i, i = 1, \ldots, n) = \prod_{i=1}^{n} c\left(u_i, v_i | \alpha_i = \exp(\beta_1 b_1(x_i) + \cdots + \beta_p b_p(x_i)\right).$$

$$\begin{aligned} & l(\beta_1, \ldots, \beta_p | u_i, v_i, i = 1, \ldots, n) \\ =& \sum_{i=1}^{n} \left\{\log \psi''(\psi^{-1}(u_i) + \psi^{-1}(v_i)) - \log[\psi'(\psi^{-1}(u_i))] - \log[\psi'(\psi^{-1}(v_i))]\right\}. \end{aligned}$$

We assigned one knot at the median of the "duration" for the natural cubic spline. Then we can fit a curve between the dependence parameter α and "duration" as in Figure 4. There are 100 gray lines used to indicate how variant the estimated α-line is. The gray lines were generated based on the simulated coefficients for the natural cubic spline, and the values of them were randomly generated by a multivariate Normal distribution with the mean being the MLEs of the $\beta_i's$, and the covariance matrix being the inverse of the Hessian matrix.

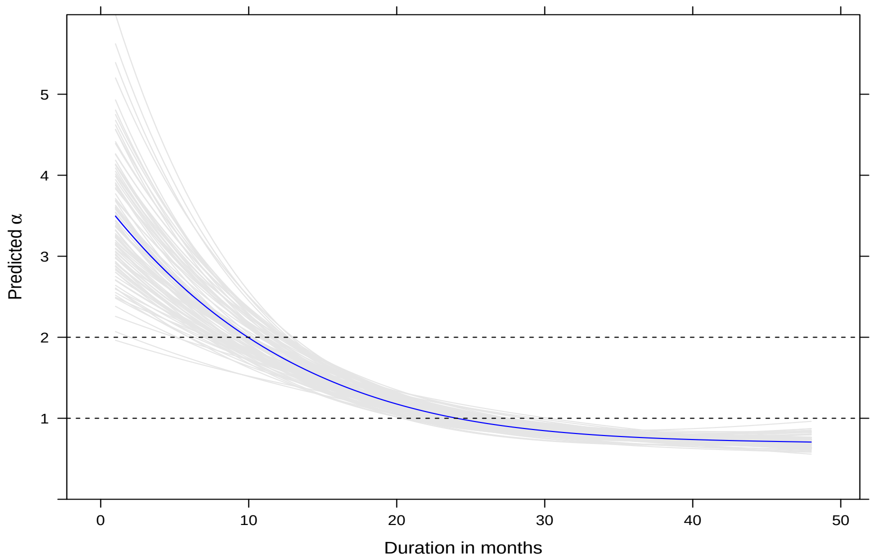
Based on Figure 4, we find that when the duration is small (≤ 10 months), the dependence in the upper tail is closed to independence ($\alpha \geq 2$); when the duration is large (≥ 24 months), the upper tail appears to be of the usual tail dependence; between 10 months and 24 months, intermediate tail dependence exists. This pattern cannot be captured by using a commonly-used copula, say, the Gumbel copula.

5 Empirical study

In Section 4, a simulation study was conducted to show that the full-range tail dependence copula can be used to account for the dynamic tail dependence patterns. In this section, we are going to apply the method for a Medical Expenditure Panel Survey data of the United States (MEPS), where both marginals and dependence are linked to covariates.

The MEPS data to be analyzed is a subset of the 2010 Full Year Consolidated Data.

Figure 4: Upper tail dependence becomes stronger as duration increases



The bivariate response variables are for emergency room visits, including expense associated with separately billing doctors (ERDEXP10: Y_1) and expense associated with facilities (ERFEXP10: Y_2). We consider covariates such as age (AGELAST) and insurance coverage (INSCOV10). We exclude all the records where at least one of the response variables is zero. Alternative methods for handling zero-inflated response variables include a logistic regression on the occurrence of zeros, and zero-inflated models such as zero-inflated Poisson and zero-inflated negative binomial regression. The sample size we are using is $2,381$. Descriptive statistics about the variables involved are in Table 2. A normalized scatter plot of the data is given in Figure 5, where each margin is transformed to be distributed as the standard Normal distribution; that is, to transform each pair of observation $(y_{1i}, y_{2i}), i = 1, \ldots, 2381$, to $(\Phi^{-1}(\text{rank}(y_{1i})/2381.5), \Phi^{-1}(\text{rank}(y_{2i})/2381.5)$, where Φ^{-1} is the inverse of the cdf of the standard Normal distribution. Based on Figure 5, we find that the dependence in the upper tail is stronger than that in the lower tail.

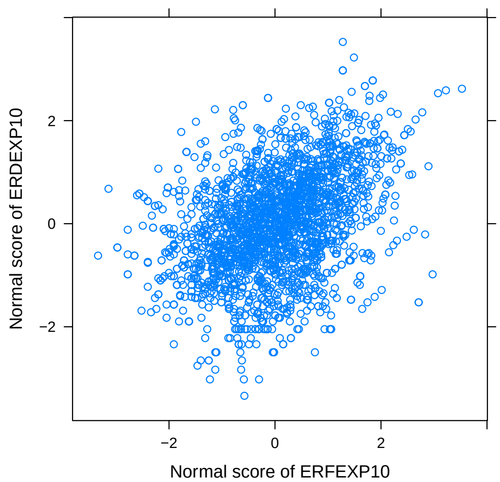
Table 2: Summary of the variables

	Min	1st quantile	Median	Mean	3rd quantile	Max
ERDEXP10	1	65	143	294.5	306	7579
ERFEXP10	1	222	520	1255	1245	50900
AGELAST	0	17	36	37.87	58	85

INSCOV10 (#obs)	Any private (1): 1126	Only public (2): 1064	Uninsured (3): 191

We first fit the two univariate response variables separately using a univariate regression.

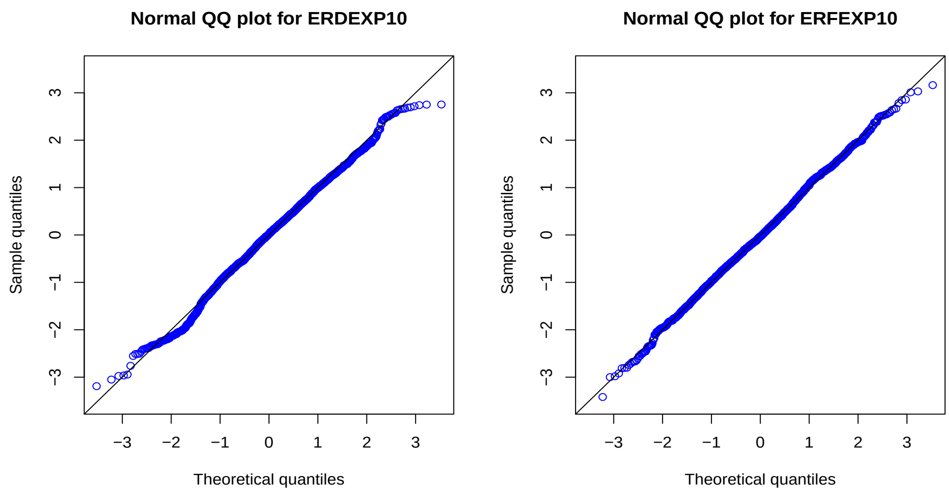
Figure 5: Normalized scatter plot for ERDEXP10 and ERFEXP10



Among many different univariate distributions, the generalized student t distribution with location and scale parameters fits the log-transformed data very well. If the random variable X follows a standard student t distribution with ν the degree of freedom, then $Y := \mu + \sigma X$ follows the generalized student t distribution with μ the location parameter and σ the scale parameter. Note that, the variance of Y is then $\sigma^2(\nu/(\nu-2))$ when $\nu > 2$. We choose some significant covariates to fit the data and their maximum likelihood estimates are reported in Table 3, and the standard errors are obtained from the observed information matrix (the inverse of the Hessian matrix). For univariate regression analysis, one can use an R package `gamlss` (Rigby and Stasinopoulos, 2005), which implements many univariate distributions including the generalized student t distribution. The Normalized QQ plots of quantile residuals are applied for diagnosis of each marginal regression (see Figure 6). We refer to Dunn and Smyth (1996) for detailed procedures on implementing the QQ plots for quantile residuals. By Dunn and Smyth (1996), when the parameters are consistently estimated, the quantile residuals converge to the standard normal distribution. The QQ plots in Figure 6 suggest that the generalized student t distribution fits the data very well. Note that, in the following, all estimates and model comparisons reported are for response variables that are transformed by the natural logarithm.

After fitting the univariate models, one can use the estimates of the coefficients of the univariate regression models as the initial values for the coefficients used in the copula regression models. Then, the inference functions for margins (IFM) method can be used to estimate the coefficients associated with the covariates that are linked to the dependence parameter. That is, transform the response variables respectively into their corresponding probabilities derived from the fitted marginal cumulative distribution functions, and then fit a copula for

Figure 6: Normalized QQ plots of quantile residuals for generalized student t regression on marginals



those transformed data. This method is fast, and can provide good initial values for the dependence parameter when maximizing a full likelihood of both marginals and dependence structures. For more details about the IFM method, we refer to Joe (1997) and the reference therein. We use the IFM method to locate initial values for the maximization, and a full likelihood method is then used to estimate the parameters for marginals and for the copulas simultaneously. The scale parameters obtained from the marginal regression are kept for the copula regression models, since the scale parameters should not affect the copula dependence structures, and this also makes computation faster and more stable. In practice, one can also consider using different location parameters than those obtained from marginal regressions.

To demonstrate how covariates can be linked to the dependence parameter, we use natural cubic splines of the "age" variable to allow flexible relationships between "age" and the dependence parameter. We choose 33.3% and 66.7% percentiles of the age variable as two knots for the natural cubic spline, and there are four coefficients $\beta_i, i = 1, 2, 3, 4$ to be estimated which are used to approximate the relationship between the covariate "age" and the dependence parameter of the copula. The maximum likelihood of the coeficients are reported in Table 3, and the standard errors (s.e.) are derived from the inverse of the Hessian matrix.

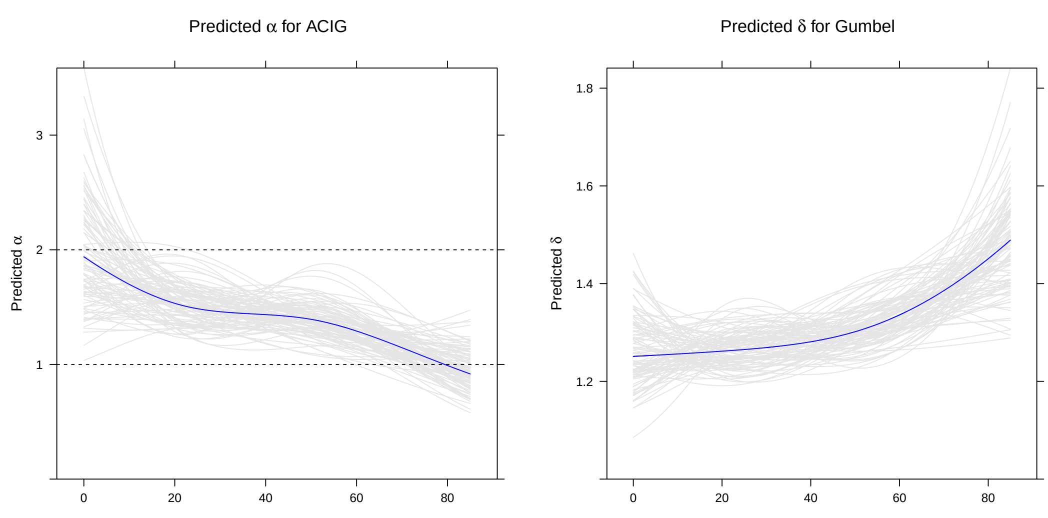
Figure 7 is the estimated relationship between age and the dependence parameters for ACIG and Gumbel models, respectively. Based on the plots, the relationship is nonlinear, and both ACIG and Gumbel copulas suggest that the dependence between expense associated with separately billing doctors and expense of facilities become stronger as one gets older.

In order to compare the ACIG regression and the Gumbel regression models, we can

Table 3: Estimates for ACIG, Gumbel and marginal regression models

		ACIG	s.e.	Gumbel	s.e.	Marginal	s.e.
ERD	(Intercept)	5.004	0.050	5.000	0.049	5.000	0.050
	age	0.007	0.001	0.007	0.001	0.007	0.001
	ins2	-0.595	0.047	-0.585	0.047	-0.589	0.048
	ins3	-0.248	0.093	-0.241	0.090	-0.260	0.091
	$\ln(\nu)$	1.847	0.093	1.878	0.092	1.919	0.146
	$\ln(\sigma)$	-0.008	–	-0.008	–	-0.008	0.026
ERF	(Intercept)	6.152	0.054	6.155	0.054	6.144	0.054
	age	0.012	0.001	0.012	0.001	0.012	0.001
	ins2	-0.703	0.052	-0.699	0.051	-0.679	0.052
	ins3	-0.123	0.098	-0.136	0.097	-0.099	0.097
	$\ln(\nu)$	2.572	0.167	2.516	0.155	2.716	0.272
	$\ln(\sigma)$	0.129	–	0.129	–	0.129	0.023
Dependence	β_1	0.226	0.241	-1.122	0.263	–	–
	β_2	0.160	0.221	-0.657	0.237	–	–
	β_3	0.813	0.263	-2.569	0.258	–	–
	β_4	-0.557	0.247	0.369	0.247	–	–

Figure 7: Upper tail dependence is fluctuated: both ACIG and Gumbel capture the dynamic structure, and ACIG can even capture the intermediate tail dependence cases.



apply the Vuong test (Vuong, 1989) for these two unnested models. The test statistic is the following

$$Z := \frac{\sqrt{n} \times \overline{m}}{\sqrt{\sum_{i=1}^{n}(m_i - \overline{m})^2}},$$

where n is the sample size, $m_i = l_i^{\mathrm{A}} - l_i^{\mathrm{G}}, i = 1, \ldots, n$, the difference between the pointwise loglikelihood of the ACIG and Gumbel models, and $\overline{m} = \frac{1}{n}\sum_{i=1}^{n} m_i$. The test statistic Z is asymptotically distributed as a standard Normal distribution when $n \to \infty$. For this dataset, $Z = -0.086$, which suggests that there is no significant difference between ACIG and Gumbel models regarding the overall fitting; these two models are quite similar in the sense of the Vuong test. However, for assessing high-risk scenarios where risks occur in the upper tails of the distributions, a criteria for the overall fitting may not be suitable in suggesting a better model. To this end, one criteria we can use is the mean square errors (MSE) beyond a certain percentile. If $\hat{y}_i, i = 1, \ldots, n$ are the predicted values and $y_i, i = 1, \ldots, n$ are the observed values, an MSE beyond a p-percentile ($p \in [0, 1]$) can be defined as

$$\mathrm{MSE}_p = \frac{1}{|I_p|}\sum_{i \in I_p}(\hat{y}_i - y_i)^2,$$

where $I_p := \{i \in \{1, \ldots, n\} : y_i > \mathrm{VaR}_p(y_i, i = 1, \ldots, n), \text{the } p\text{-percentile of the } y_i\text{'s}\}$, and $|I_p|$ is the number of elements in the index set I_p. The value of MSE_p suggests how well a model fits the tail beyond the p-percentile.

For this dataset, although the Vuong's test cannot suggest a better model, we can use the MSE_p to compare the models and conclude that the ACIG regression model can better model the upper tail and lead to a lower MSE_p as p is sufficiently large. The results are reported in Table 4, from which we find that: (1) the ACIG model is generally better than the Gumbel model in assessing the high-risk scenarios, and the advantage becomes more clear when p is relatively larger; (2) regression models with ACIG can improve the fitting of upper tails for each marginals, which justifies the benefit of combing two univariate regression models using an appropriately chosen copula; (3) the advantage of ACIG becomes relatively more significant when assessing the aggregated dependent risks.

Note that, according to MPE_p, the improvement of the Gumbel model over the independent model is small. This suggests that the assessment of the sum of the two response variables in the dataset is not very sensitive to the dependence structures between the response variables. It may happen because the assessment depends not only on the dependence structure but also on the marginals. That said, the improvement of the ACIG model over the Gumbel model is relatively more significant, although the difference of mean square errors is not large in absolute values. The improvement in modeling the upper tail is probably due to the full-range tail dependence property of the ACIG copula, so that the upper tail can be better taken care of. Since the two response variables are dependent on each other, and their dependence relations may not be completely explained by available covariates, a proper modeling of their dependence structure can help better explain the behavior of each individual response variable as well.

Table 4: MSE_p for ACIG, Gumbel and marginal models. The MSE_p for the ACIG model is generally lower, especially for assessing the sum of the expenses.

	p	0.1	0.3	0.5	0.7	0.9	0.95	0.99	0.995	0.999
	# obs	2142	1666	1190	714	238	118	24	12	3
ERD	ACIG	0.92	0.93	1.23	1.93	4.12	6.00	11.98	13.55	14.27
	Gumbel	0.93	0.94	1.24	1.96	4.16	6.04	12.05	13.63	14.35
	Indep.	0.93	0.95	1.25	1.97	4.17	6.06	12.10	13.68	14.35
ERF	ACIG	1.13	1.15	1.48	2.30	4.57	6.36	10.86	12.66	17.60
	Gumbel	1.13	1.15	1.49	2.31	4.59	6.39	10.91	12.71	17.70
	Indep.	1.13	1.16	1.49	2.33	4.63	6.43	10.93	12.74	17.65
ERD	ACIG	3.08	3.05	3.88	5.99	12.59	17.80	33.56	40.76	57.56
+	Gumbel	3.08	3.07	3.92	6.05	12.72	17.95	33.78	40.98	57.92
ERF	Indep.	3.08	3.07	3.94	6.09	12.79	18.04	33.86	41.07	58.00

We can use the QQ plot of quantile residuals (Dunn and Smyth (1996)) to evaluate the fitting of the quantity of interest. We now consider the QQ plot for the sum of ERDEXP10 and ERFEXP10. Let $S \overset{d}{=} Y_1+Y_2$, where Y_1 and Y_2 are continuous random variables supported on $(-\infty, \infty)$. Let F be the cdf of S, F_1 and F_2 be the cdfs for Y_1 and Y_2, respectively, and $C(u_1, u_2)$ be the copula for Y_1 and Y_2. Then

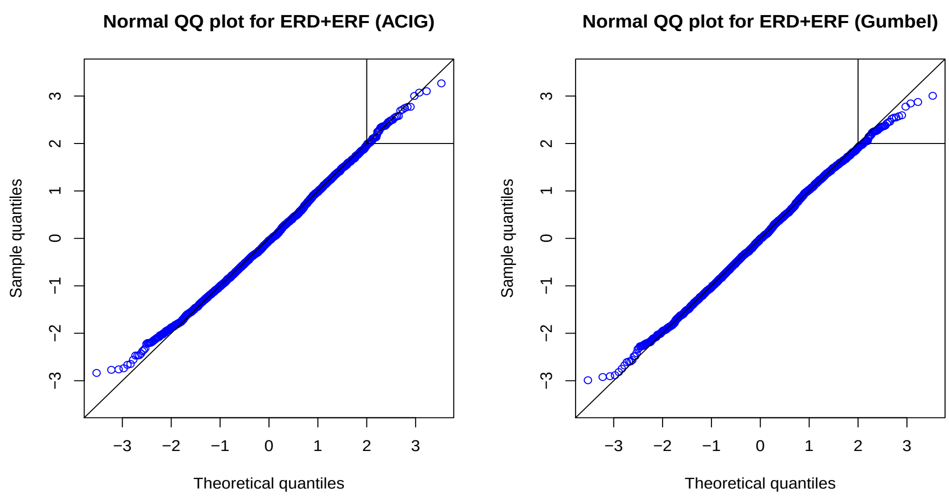
$$F(s) = \int_0^1 C_{1|2}\left(F_1(s - F_2^{-1}(u)), u\right) du \qquad (6)$$

where $C_{1|2}(u_1, u_2) = \partial C(u_1, u_2)/\partial u_2 = \psi'(\psi^{-1}(u_1) + \psi^{-1}(u_2))/\psi'(\psi^{-1}(u_2))$.

Based on (6) and the estimates of the parameters for the copula and the marginals, we can evaluate the estimated $\hat{F}(s)$. Then, for each pair of observed response variables, we can calculate the sum of the two values, denoted as $s_i, i = 1, \ldots, n$, where n is the sample size. Then plug s_i's into $z_i = \Phi^{-1}(\hat{F}(s_i))$ to get z_i's. Plotting z_i's against the standard Normal distribution leads to the normalized QQ plot of quantile residuals for the sum of Y_1 and Y_2. For a regression setting, the cdf F, copula C and the marginal distributions in (6) can be simply replaced by their conditional versions that are conditioning on the values of covariates, and the QQ plot still works. We refer to Dunn and Smyth (1996) for more details. The integration in (6) can be calculated numerically.

The normalized QQ plots of the quantile residuals for the sum of ERDEXP10 and ERFEXP10 are in Figure 8, based on which one can find that: (1) the overall performance of ACIG and Gumbel models are similar; (2) the regression model based on the ACIG copula outperforms in the upper tail where risks occur, while the Gumbel model leads to a heavier right tail of the sum, thus overestimating the losses; (3) the cost of the better fitting in the upper tail for the ACIG model is the relatively worse fitting in the lower tail; however the lower tail is not important for the purpose of risk assessment.

Figure 8: Normalized QQ plots of quantile residuals for ERD+ERF



6 Concluding remark

The research work was motivated by observing dynamic dependence structures between loss and ALAE of auto bodily injury insurance claim datasets. When covariates are available, one can use the covariates to explain not only the marginals but also the dependence structure itself. Because the dependence in the upper tail can be relatively weaker or stronger conditioning on different values of covariates, it is better to use a copula that has the widest range of dependence in the upper tail to model the dependence structure between these two response variables, and let the data explain how strong the dependence is. The commonly-used Gumbel copula only has the usual upper tail dependence. In other words, no matter what the degree of dependence in the upper tail of the data is, a fitted Gumbel copula always suggests a relatively stronger local dependence in the upper tail. To this end, ACIG is a good candidate model because the upper tail is very flexible, and the only dependence parameter of the copula simply captures the most important information of our interest – the dependence in the upper tail, which makes the difference of the overall dependence structures.

A statistical criteria based on the overall fitting may not be appropriate for comparing models used for assessing high-risk scenarios. For the empirical example, the regression model based on the ACIG copula, although does not outperform in overall fitting, has better fitting for the upper tail, especially when it is used for assessing the tail risk of aggregated dependent losses.

Dependence between the response variables may be well explained by the covariates through their effects on the marginals. When the covariates cannot explain the dependence between response variables completely through their effects on the marginals, a regression on dependence parameters could become useful. Moreover, a strong tail dependence structure that can be observed marginally without conditioning on any covariates may still need a relatively weaker tail dependence copula to model when covariates are available. For example, the aforementioned well known dataset of loss and ALAE studied in the actuarial literature appears to have relatively stronger upper tail dependence. However, if covariates are available for the dataset, then the relatively stronger upper tail dependence could be partially explained by some common effects of the covariates on the marginals, and then a weaker tail dependence copula could be necessary. To this end, a full-range tail dependence copula such as the ACIG copula becomes useful. Depending on how much tail dependence that can be partially explained by covariates, the copula used to account for the dependence in the tails should be flexible enough to capture the dynamic tail dependence structures, which could be very weak or very strong.

The regression model can also be extended to the case where response variables have more than two dimensions. Then a critical task is to model the dependence between the response variables appropriately, which is often challenging when the dimension becomes large. Dependence modeling for high-dimensional random variables are nowadays under active development. To name a few of the techniques, we refer the interested reader to Vine copula (Bedford and Cooke, 2002; Kurowicka and Joe, 2011), factor copula (Krupskii and Joe, 2013; Oh and Patton, 2012), and copulas implemented in probabilistic graphical models (Elidan, 2013).

7 Details about numerical issues

When implementing a new statistical model, the computation speed is often important. The R functions (with C routines) for the regression models with ACIG copula and with Gumbel copula have been implemented. After implementing the ACIG copula using C codes, the speed of calculations for the ACIG copula becomes much faster, and the overall performance is satisfactory for dealing with real applications.

In what follows, we present some numerical details that were considered when implementing the models. In order to improve the speed for functions used for the ACIG copula, C codes have been used for those time-consuming calculations.

7.1 ACIG copula

1. To get the inverse of the LT of inverse Gamma, we take a natural logarithm for both sides of $\psi(s) = t$ for $0 < t \le 1$, and then use a Newton's method to find the root. The reason is that a logarithm of a gamma function can handle large arguments while the gamma function itself often has numerical errors for large arguments.

$$\psi(s;\alpha) = \frac{2}{\Gamma(\alpha)} s^{\alpha/2} K_\alpha(2\sqrt{s}), \quad s \ge 0,\ \alpha > 0.$$

   Therefore, it remains to solve

$$g(s) := \frac{\alpha}{2}\ln(s) + \ln(K_\alpha(2\sqrt{s})) - \ln(t) - \ln(\Gamma(\alpha)) + \ln(2) = 0.$$

   Also,

$$\begin{aligned} g'(s) &:= \frac{\alpha}{2s} - \frac{K_{\alpha-1}(2\sqrt{s}) + K_{\alpha+1}(2\sqrt{s})}{2\sqrt{s}K_\alpha(2\sqrt{s})} \\ &= \frac{\alpha}{2s} - \exp\{\ln(K_{\alpha-1}(2\sqrt{s})) - \ln(2\sqrt{s}) - \ln(K_\alpha(2\sqrt{s}))\} \\ &\qquad - \exp\{\ln(K_{\alpha+1}(2\sqrt{s})) - \ln(2\sqrt{s}) - \ln(K_\alpha(2\sqrt{s}))\}. \end{aligned}$$

   Here the functions $\ln K_\alpha()$ and $\ln\Gamma()$ can handle relatively larger arguments and are more stable numerically.

2. For the joint density function of F_1 and F_2 where the dependence is modeled by the ACIG copula C, since

$$\begin{aligned} \psi'(s) &= -2s^{(\alpha-1)/2}K_{\alpha-1}(2\sqrt{s})/\Gamma(\alpha); \\ \psi''(s) &= 2s^{(\alpha-2)/2}K_{\alpha-2}(2\sqrt{s})/\Gamma(\alpha), \end{aligned}$$

   letting $s := \psi^{-1}(F_1(x_1)) + \psi^{-1}(F_2(x_2))$, $s_1 := \psi^{-1}(F_1(x_1))$ and $s_2 := \psi^{-1}(F_2(x_2))$, we have

$$\begin{aligned}\ln f(x_1,x_2) &= \ln\psi''(s)+\ln f_1(x_1)+\ln f_2(x_2)-\ln(-\psi'(s_1))-\ln(-\psi'(s_2))\\ &= \ln K_{\alpha-2}(2\sqrt{s})-\ln K_{\alpha-1}(2\sqrt{s_1})-\ln K_{\alpha-1}(2\sqrt{s_2})\\ &\quad+\frac{\alpha-2}{2}\ln(s)-\frac{\alpha-1}{2}[\ln(s_1)+\ln(s_2)]\\ &\quad+\ln f_1(x_1)+\ln f_2(x_2)-\ln(2)+\ln\Gamma(\alpha),\end{aligned}$$

where $\ln K_\alpha()$ and $\ln\Gamma()$ can handle relatively larger arguments and are more stable numerically.

3. For $C_{1|2}(u_1,u_2)$, assuming $s_1=\psi^{-1}(u_1)$, $s_2=\psi^{-1}(u_2)$, and $s=s_1+s_2$, we obtain its logarithm as

$$\begin{aligned}\log C_{1|2}(u_1,u_2) &= \log(-\psi'(s))-\log(-\psi'(s_2))\\ &= \frac{\alpha-1}{2}(\log(s)-\log(s_2))+\log K_{\alpha-1}(2\sqrt{s})-\log K_{\alpha-1}(2\sqrt{s_2})\end{aligned}$$

7.2 Gumbel copula

1. For the Gumbel copula, the LT ψ has the following properties.

$$\begin{aligned}\psi(s) &= \exp\{-s^{1/\theta}\},\quad \theta\ge 1;\\ \psi'(s) &= -\exp\{-s^{1/\theta}\}s^{1/\theta-1}/\theta,\quad \theta\ge 1;\\ \psi''(s) &= \exp\{-s^{1/\theta}\}\left(s^{1/\theta-1}/\theta\right)^2-\exp\{-s^{1/\theta}\}(1/\theta-1)s^{1/\theta-2}/\theta,\quad \theta\ge 1;\\ \psi^{-1}(s) &= (-\ln(s))^\theta,\quad \theta\ge 1.\end{aligned}$$

Then, letting $s:=\psi^{-1}(F_1(x_1))+\psi^{-1}(F_2(x_2))$, $s_1:=\psi^{-1}(F_1(x_1))$ and $s_2:=\psi^{-1}(F_2(x_2))$, we have

$$\begin{aligned}\ln f(x_1,x_2) &= \ln\psi''(s)+\ln f_1(x_1)+\ln f_2(x_2)-\ln(-\psi'(s_1))-\ln(-\psi'(s_2))\\ &= -s^{1/\theta}+\ln\left[s^{2/\theta-2}/\theta^2-(1/\theta^2-1/\theta)s^{1/\theta-2}\right]+\ln f_1(x_1)+\ln f_2(x_2)\\ &\quad-\ln(F_1(x_1))+2\ln(\theta)-(1-\theta)[\ln(-\ln(F_1(x_1)))]\\ &\quad-\ln(F_2(x_2))-(1-\theta)[\ln(-\ln(F_2(x_2)))].\end{aligned}$$

2. For $C_{1|2}(u_1,u_2)$, assuming $s_1=\psi^{-1}(u_1)$, $s_2=\psi^{-1}(u_2)$, and $s=s_1+s_2$, we obtain the

natural logarithm as

$$\log C_{1|2}(u_1, u_2) = \log(-\psi'(s)) - \log(-\psi'(s_2)) = (1 - 1/\theta)[\log(s_2) - \log(s)] + s_2^{1/\theta} - s^{1/\theta}.$$

References

Acar, E., Craiu, R., and Yao, F. (2011). Dependence calibration in conditional copulas: A nonparametric approach. *Biometrics*, 67:445–453.

Bedford, T. and Cooke, R. M. (2002). Vines—a new graphical model for dependent random variables. *The Annals of Statistics*, 30(4):1031–1068.

Czado, C., Kastenmeier, R., Brechmann, E., and Min, A. (2011). A Mixed Copula Model for Insurance Claims and Claim Sizes. *Scand. Actuar. J.*

de Leon, A. and Wu, B. (2011). Copula-based regression models for a bivariate mixed discrete and continuous outcome. *Statistics in Medicine*, 30:175–185.

Donnelly, C. and Embrechts, P. (2010). The devil is in the tails: actuarial mathematics and the subprime mortgage crisis. *Astin Bulletin*, 40(1):1–33.

Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5(3):236–244.

Elidan, G. (2013). Copulas in machine learning. In *Copulae in Mathematical and Quantitative Finance*, pages 39–60. Springer.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal*, 2(1):1–25.

Hua, L. (2013). Tail negative dependence and its applications for aggregate loss modeling. *Technical Report*.

Hua, L. and Joe, H. (2011). Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis*, 102:1454–1471.

Hua, L. and Joe, H. (2013). Intermediate tail dependence: a review and some new results. In Li, H. and Li, X., editors, *Stochastic Orders in Reliability and Risk: In Honor of Professor Moshe Shaked*, chapter 15, pages 291–311. Springer.

Hua, L. and Joe, H. (2014). Strength of tail dependence based on conditional tail expectation. *Journal of Multivariate Analysis*, pages 143–159.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Klugman, S. A. and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics*, 24(1-2):139–148. 1st IME Conference (Amsterdam, 1997).

Krupskii, P. and Joe, H. (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis*, 120:85 – 101.

Kurowicka, D. and Joe, H. (2011). *Dependence Modeling: Vine Copula Handbook*. World Scientific Publishing Company, Singapore.

Oh, D. and Patton, A. (2012). Modelling dependence in high dimensions with factor copulas. *Technical report, Duke University.*

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape,(with discussion). *Applied Statistics*, 54:507–554.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, pages 307–333.