ON BIVARIATE EXTREME VALUE COPULAS IN ACTUARIAL SCIENCE

Werner Hürlimann

Feldstrasse 145, CH-8004 Zürich

ABSTRACT

The dependence structure between largest claims of two different lines of insurance business is analyzed using the bivariate extreme-value copula. We use non-parametric inference of the dependence function with parametric estimation of the marginal distributions to evaluate a statistical model for the purpose of actuarial calculations. Original applications to return periods of bivariate events, to the retained risk of a bivariate portfolio and to the bivariate rate on line for excess-of-loss reinsurance layers are given. Illustration is based on bivariate losses from a reinsurance captive. The available scarce data is typical for this insurance business, which can only be analysed with limited statistical inference methodology.

Keywords: *bivariate extreme-value copula, asymmetric logistic model, Gumbel model, bivariate return period, retained risk, bivariate rate on line, excess-of-loss reinsurance.*

1. INTRODUCTION

It is well-known that large claims are predominant in the analysis of non-life insurance losses, especially in the context of excess-of-loss reinsurance. While the univariate analysis of large claims data is often discussed in the literature, the actuarial study of dependent large claims is more recent (e.g. Cebrian et al. [1]).

The most general margin-free way to describe the dependence structure of a multivariate distribution is through its copula function introduced in Sklar [2]. A very suitable way to analyze dependent large claims data is through extreme-value copulas. They arise as the possible limits of copulas of component-wise maximum of independent and identically distributed samples. For a bivariate sample $(X_{i1}, X_{i2}), i = 1,..., n,$ the vector of component-wise maxima is (M_{n1}, M_{n2}) with $M_{ij} = \max_{1\le i\le n}\{X_{ij}\}, j = 1,2$. If the pairs (X_{i1}, X_{i2}) are independent and if they have a common bivariate distribution function F with continuous margins and copula C_F, then the joint distribution function of (M_{n1}, M_{n2}) is F^n with copula $C_{F^n}(u_1,u_2) = C_F^n\left(u_1^{1/n}, u_2^{1/n}\right)$ for $(u_1,u_2)\in[0,1]^2$. If there exists a copula C such that $C_{F^n}(u_1,u_2) \to C(u_1,u_2)$ as $n\to\infty$, then C is called *bivariate extreme-value* (BEV) *copula*.

In the actuarial context, if the pair (X_{i1}, X_{i2}) records the claims of an insurance business over a given period for two different lines of business, then the pair (M_{n1}, M_{n2}) represents the maximal claims recorded during a given period. If (x_1, x_2) are given claim sizes of the two lines of business, then $F^n(x_1, x_2)$ is the probability that the largest claims components will be below the components of (x_1, x_2) over a given period such that $M_{n1} \le x_1, M_{n2} \le x_2$. The copula C_{F^n} of F^n describes the dependence between the occurrences of claims for the two lines of business. As n is large, asymptotic theory suggests to model C_{F^n} by an extreme-value copula. A copula $C(u,v)$ is a BEV copula if and only if there exists a real-valued function $A(t)$ on the interval $[0,1]$ such that $C(u,v) = \exp\left\{\ln(uv)\cdot A\left(\frac{\ln(v)}{\ln(uv)}\right)\right\}$ for $0 < u,v < 1$, where the dependence function $A(t)$ satisfies well-known constraints.

The present contribution is structured as follows. Information on BEV copulas are found in Section 2 and related literature. Section 3 explains how the return periods of random events depending upon the joint behavior of two dependent random variables may be evaluated in general and for BEV copulas in particular. This is of importance in the practice of reinsurance. If (X,Y) represents a random pair of claims from two dependent lines of business, then it is essential to assess the return periods of "dangerous" events. For example, one is interested in the event where X or Y (alternatively X and Y) exceeds given thresholds x and y, that is in the return period for the event $\{X > x\} \vee \{Y > y\}$ (alternatively $\{X > x\} \wedge \{Y > y\}$). Proposition 3.1 is an original and new characterization of return periods, which only depends on the Pickands dependence function of a bivariate extreme-value copula. A previous alternative analysis of return periods in a bivariate context is found in Hürlimann [3]. Section 4 presents two important formulas required to evaluate standard deviation premiums for the retained risk on a bivariate portfolio when the cedent transfers risk through an excess-of-loss reinsurance contract. The formulas obtained in Section 4 are further used in Section 5 to assess a newly defined actuarial notion of bivariate rate on line for excess-of-loss reinsurance layers. A short account of available statistical inference methodology for a BEV copula is summarized in Section 6 and further discussed in Section 7. Finally, Section 7 illustrates numerically all of the discussed topics using a small sample data set from insurance practice, which collects largest claims pairs of insurance policies over a given period for product liability (PL) and general liability (GL) insurance. The one-parameter Gumbel copula with Pareto lognormal margins turns out to be our preferred bivariate model in the context of risk management for reinsurance captives of corporate firms, for which typically only a few observations are available for actuarial evaluation.

2. BIVARIATE EXTREME-VALUE COPULAS

The dependence structure between large claims of different lines of insurance business is analyzed using bivariate extreme value copulas or BEV copulas. According to Pickands [4] (see also Tawn [5]) a copula $C(u,v)$ is a BEV copula if and only if there exists a real-valued function $A(t)$ on the interval $[0,1]$ such that

$$C(u,v) = \exp\left\{ \ln(uv) \cdot A\left(\frac{\ln(u)}{\ln(uv)} \right) \right\}, \quad 0 < u,v < 1. \tag{1}$$

The so-called *Pickands dependence function* $A(t)$ must satisfy the following two properties:

(i) $\max(t, 1-t) \le A(t) \le 1$

(ii) $A(t)$ is convex

In particular, the endpoint conditions $A(0) = A(1) = 1$ must hold. The upper bound $A(t) \equiv 1$ corresponds to the independent copula $C(u,v) = uv$ and the lower bound $A(t) = \max(t, 1-t)$ to the comonotone copula $C(u,v) = \min(u,v)$.

The BEV copula satisfies some important stochastic ordering properties. In particular, it is monotone regression dependent in the terminology of Lehmann [6], or in modern terms, it satisfies the stochastic increasing property. That is, the conditional distribution $P(Y > y | X = x)$ is non-decreasing in x and the conditional distribution $P(X > x | Y = y)$ is non-decreasing in y (e.g. Garralda Guillem [7]). For stochastically increasing distributions, a conjecture by Hutchinson and Lai [8] states that $-1 + \sqrt{1+3\tau} \le \rho_S \le \min\left\{\frac{3}{2}\tau, 2\tau - \tau^2\right\}$, where τ and ρ_S denote Kendall's tau and Spearman's rho, respectively. As shown in Hürlimann [9], the conjecture holds for the class of bivariate extreme-value copulas. The interested reader may found extended results to the Archimax class of copulas in Hürlimann [10].

Several parametric models for the dependence function are available in the literature (see e.g. Kotz and Nadarajah [11] and Beirlant et al. [12]). For illustration we retain the most simple and popular Gumbel model. Variability of the results against changes in the model specification is analyzed in Section 7 using the extended asymmetric logistic model.

<u>Logistic model (Gumbel [13])</u>

A simple and popular parametric model is the Gumbel or logistic model $A(t)=\left[t^r+(1-t)^r\right]^{\frac{1}{r}}$ with $r\geq 1$. The corresponding copula is the only BEV copula that is also an Archimedean copula (Genest and Rivest [14]). Independence and complete dependence correspond to $r=1$ and $r=\infty$ respectively.

<u>Asymmetric logistic model</u>

The dependence function is $A(t)=\left[\theta^r\cdot(1-t)^r+\phi^r\cdot t^r\right]^{\frac{1}{r}}+(\theta-\phi)\cdot t+1-\theta$ with $\theta\geq 0$, $\phi\leq 1$ and $r\geq 1$. This model is very flexible and contains several common models like the Gumbel or logistic $(\theta=\phi=1)$, the biextremal, the dual of the biextremal, as well as a mixture of Gumbel and independence models. Independence holds for $\theta=0$ or $\phi=0$ or $r=1$ and complete dependence holds when $\theta=\phi=1$ and $r=\infty$.

Some important properties of BEV copulas can be formulated in terms of the dependence function $A(t)$. For example, Kendall's τ is given by ($A'(t)$ denotes the right derivative of $A(t)$)

$$\tau_A=\int_0^1\frac{t(1-t)}{A(t)}\cdot dA'(t)dt \tag{2}$$

<u>Gumbel model</u>

$$\tau_A=1-r^{-1}$$

<u>Asymmetric logistic model</u>

$$\tau_A=\int_0^1\frac{t(1-t)(r-1)\theta^r\phi^r\left(1-t^2\right)^{r-2}}{\left(\theta^r\cdot(1-t)^r+\phi^r\cdot t^r\right)}dt$$

3. RETURN PERIODS FOR BEV COPULAS

The *return period* of a given event is defined as the average time elapsing between two successive realizations of the event itself. The return period provides a very simple, yet efficient, means of risk analysis. It is able to summarize information about the extreme behavior of stochastic phenomena into a single number. A general copula framework for studying the return periods of random events depending upon the joint behavior of two non-independent random variables is provided in Salvadori [15]. In non-financial research the return period is a well-established concept (see Salvadori et al. [16], Shiau [17], etc.).

Consider a sequence $E_1, E_2, \ldots$ of independent events that may happen at times $t_1<t_2<\ldots$ (temporal marked point process). Each event is characterized by the joint behavior of a pair of uniform random variables (U,V) with copula $C(u,v)$, which can be expressed in terms of the marginal events $\{U\leq u\}, \{U>u\}, \{V\leq v\}, \{V>v\}$. In practice of reinsurance, an event is defined as "dangerous" if

(i) *either* U *or* V exceeds given thresholds, in which case the event $E_{u,v}^{\vee}=\{U>u\}\vee\{V>v\}$ is relevant.

(ii) U *and* V are larger than prescribed values, in which case the event $E_{u,v}^{\wedge}=\{U>u\}\wedge\{V>v\}$ is relevant.

Now, let T_i be the *inter-arrival time* between E_i and E_{i+1}, and assume that the mean $E[T_i]$ exists and is finite. Let $N_{u,v}^{\vee}$ and $N_{u,v}^{\wedge}$ denote, respectively, the number of events E_i between two successive realizations of $E_{u,v}^{\vee}$ and $E_{u,v}^{\wedge}$, and let $T_{u,v}^{\vee}$ and $T_{u,v}^{\wedge}$ be, respectively, the inter-arrival time between two successive realizations of $E_{u,v}^{\vee}$ and $E_{u,v}^{\wedge}$. It turns out that

$$T_{u,v}^{\vee}=\sum_{i=1}^{N_{u,v}^{\vee}}T_i,\quad T_{u,v}^{\wedge}=\sum_{i=1}^{N_{u,v}^{\wedge}}T_i\,. \tag{3}$$

Under the assumption that the T_i's are independent and identically distributed (and independent of U and V) it follows that

$$\begin{aligned} \tau_{u,v}^{\vee} &= E[T_{u,v}^{\vee}] = E[N_{u,v}^{\vee}]\cdot E[T], \\ \tau_{u,v}^{\wedge} &= E[T_{u,v}^{\wedge}] = E[N_{u,v}^{\wedge}]\cdot E[T], \end{aligned} \tag{4}$$

where T denotes any of the random variables T_i. Clearly, $N_{u,v}^{\vee}$ and $N_{u,v}^{\wedge}$ have a geometric distribution with parameters $p_{u,v}^{\vee}$ and $p_{u,v}^{\wedge}$ given by, respectively

$$\begin{aligned} p_{u,v}^{\vee} &= P(U > u \vee V > v) = 1 - C(u,v), \\ p_{u,v}^{\wedge} &= P(U > u \wedge V > v) = (1-u) + (1-v) - [1 - C(u,v)]. \end{aligned} \tag{5}$$

From (4) it follows that

$$\tau_{u,v}^{\vee} = \frac{E[T]}{p_{u,v}^{\vee}}, \quad \tau_{u,v}^{\wedge} = \frac{E[T]}{p_{u,v}^{\wedge}}. \tag{6}$$

These numbers represent the return periods of the events $E_{u,v}^{\vee}$ and $E_{u,v}^{\wedge}$. Now, for a pair of arbitrary random variables (X,Y) with marginal distributions $F_X(x), F_Y(y)$ and copula $C(u,v)$, the return periods of the events $E_{x,y}^{\vee} = \{X > x\} \vee \{Y > y\}$ and $E_{x,y}^{\wedge} = \{X > x\} \wedge \{Y > y\}$ are given by

$$\tau_{x,y}^{\vee} = \frac{1}{f\cdot[1 - C[F_X(x), F_Y(y)]]}, \quad \tau_{x,y}^{\wedge} = \frac{1}{f\cdot[\overline{F}_X(x) + \overline{F}_Y(y) - [1 - C[F_X(x), F_Y(y)]]]}. \tag{7}$$

where $f = E[T]^{-1}$ denotes the frequency at which the events E_i happen.

For a BEV copula, it is possible to characterize the *isolines of constant return period* in terms of the dependence function $A(t)$. The next result states this for the event $E_{x,y}^{\vee}$.

Proposition 3.1. Let (X,Y) be a random pair with absolutely continuous margins $F_X(x), F_Y(y)$ and BEV copula (1), and let f be the frequency of events E_i between two successive realizations of the event $E_{x,y}^{\vee} = \{X > x\} \vee \{Y > y\}$. The isolines of constant return period $\tau_{x,y}^{\vee} = T$ are characterized by the marginal probabilities

$$F_X(x) = \left(1 - \frac{1}{f\cdot T}\right)^{\frac{\xi}{A(\xi)}}, \quad F_Y(y) = \left(1 - \frac{1}{f\cdot T}\right)^{\frac{1-\xi}{A(\xi)}}, \quad \xi = \frac{\ln F_Y(x)}{\ln F_X(x) + \ln F_Y(y)}. \tag{8}$$

Proof. It suffices to parameterize the level curves of the BEV copula by solving the equation $C(u,v) = t$, that is

$$\ln(uv)\cdot A\left(\frac{\ln(u)}{\ln(uv)}\right) = \ln(t). \tag{9}$$

In a first step, the substitution $z = \ln(uv)$ allows to write the coordinates u, v as functions of the parameters z and t as follows:

$$u(z,t) = \exp\left\{z\cdot A^{-1}\left(\frac{\ln t}{z}\right)\right\}, \quad v(z,t) = \exp\left\{z\left[1 - A^{-1}\left(\frac{\ln t}{z}\right)\right]\right\}. \tag{10}$$

In a second step, setting $\xi = \frac{\ln v}{z} = A^{-1}\left(\frac{\ln t}{z}\right)$, one sees that $z = \frac{\ln t}{A(\xi)}$, which inserted in (10) yields

$$u(\xi,t) = \exp\left\{\frac{\xi}{A(\xi)} \cdot \ln t\right\} = t^{\frac{\xi}{A(\xi)}}, \quad v(\xi,t) = \exp\left\{\frac{1-\xi}{A(\xi)} \cdot \ln t\right\} = t^{\frac{1-\xi}{A(\xi)}}. \tag{11}$$

According to (7), the return period $\tau_{x,y}^{\vee}$ will be equal to T provided the marginal probabilities satisfy (8). ◊

Let us analyze somewhat the behavior of these probabilities for a constant return period. We use the fact that the function $h(\xi) = \frac{1-\xi}{A(\xi)}$ is non-increasing for all BEV copulas. This property follows from Galambos [18], equation (31), p. 261, which states that the function $(1+e^{-y})v(y)$ is non-increasing in y, where $v(y)$ is considered such that $A(\xi) = v(\ln \frac{1-\xi}{\xi})$, and the decreasing transformation $y = \ln \frac{1-\xi}{\xi}$ is used. Since $h'(\xi) < 0$ the probability function $q_Y(\xi) = F_Y(x)$ in (8) is monotone increasing from $1-(fT)^{-1}$ to 1 for all BEV copulas. Moreover, in case $A(\xi)$ is a symmetric function such that $A(\xi) = A(1-\xi)$, the probability function $q_X(\xi) = F_X(y)$ in (8) satisfies $q_X(\xi) = q_Y(1-\xi)$ and is monotone decreasing from 1 to $1-(fT)^{-1}$. For symmetric dependence functions both probability functions are equal exactly when $\xi = \frac{1}{2}$, in which case one has uniquely $q_X\left(\frac{1}{2}\right) = q_Y\left(\frac{1}{2}\right) = \left(1-(fT)^{-1}\right)^{\frac{1}{2}A\left(\frac{1}{2}\right)^{-1}}$. We observe that the dependence functions of the Gumbel model and the asymmetric logistic model with $\phi = \theta$ are both symmetric and the preceding analysis applies to them.

Remark 3.1. As possible further development, it would be interesting to consider also the “secondary return period” first introduced by Salvadori [15]. Indeed, thanks to the proper use of Kendall’s distribution or multivariate probability integral transform (PIT), the secondary return period reduces the bivariate (more generally, multivariate) dynamics to a one-dimensional one, and yields a parameter similar to the univariate return period. More information on Kendall’s distribution and its applications is found in the papers by Genest and Rivest [19], [20], Capéraà et al. [21], Ghoudi et al. [22], Chakak and Ezzerg [23], Chakak and Imlahi [24], Nelsen et al. [25], Genest and Boies [26], Genest et al. [27], [28], Kolev et al. [29], Belzunce et al. [30] and Brechmann [31] among others.

4. RETAINED RISK FOR BIVARIATE RISKS

It is well-known that the dependence structure between two portfolios of risk influences the excess-of-loss premiums for the aggregate portfolio of risks (e.g. Heilmann [32], Hürlimann [33], [34], Dhaene and Goovaerts [35]). Clearly, the same holds true for the retained risk of a cedent, which transfers risk through an excess-of-loss reinsurance contract.

Given a random pair (X,Y) with fixed margins $F_X(x), F_Y(y)$ and bivariate copula $C(u,v)$, the next result shows how *standard deviation premiums* for the retained risk in a bivariate excess-of-loss reinsurance contract can be evaluated. For a deductible d, denote the reinsurance payoff by $Z_d(X,Y) = (X+Y-d)_+$, the retained risk by $R_d(X,Y) = \min(d, X+Y) = d-(d-X-Y)_+$. It suffices to obtain formulas for the mean and second-order moment of $(d-X-Y)_+$.

Proposition 4.1. Let (X,Y) be a random pair of non-negative random variables with absolutely continuous margins $F_X(x), F_Y(y)$ and bivariate copula $C(u,v)$. Then one has the formulas

$$E[(d-X-Y)_+] = \int_0^d C[F_X(x), F_Y(d-x)]dx \tag{12}$$

$$E[(d-X-Y)_+^2] = 2 \cdot \iint_{0 \le x+y \le d} C[F_X(x), F_Y(y)]dxdy \tag{13}$$

Proof. Without copula representation, the formula (12) follows from Lemma 2 of Dhaene and Goovaerts [35]. The integral representation (13) can be shown along the same lines. ◊

5. BIVARIATE RATE ON LINE FOR EXCESS-OF-LOSS REINSURANCE LAYERS

In excess-of-loss reinsurance the premium divided by the cover is called *rate on line* (ROL). It is the inverse of the *payback period*, which represents the number of years at a given level premium that would be necessary to accumulate total premiums equal to the cover.

Let (X,Y) be a random pair representing the claim sizes of two lines of business. Suppose that λ_X,λ_Y are the expected numbers of claims with claims sizes X and Y respectively. Merging the two lines of business to a single multi-line yields a claim size defined by

$$Z=\left(\frac{\lambda_X}{\lambda}\right)X+\left(\frac{\lambda_Y}{\lambda}\right)Y, \tag{14}$$

where $\lambda=\lambda_X+\lambda_Y$. One is interested in the bivariate ROL defined by

$$ROL(d,L)=\frac{E[(Z-d)_+]-E[(Z-L)_+]}{L-d}. \tag{15}$$

Given the distributions $F_X(x), F_Y(y)$ of the marginals and a copula function $C(u,v)$ describing the bivariate dependence, the calculation of the bivariate ROL uses the equivalent formula

$$ROL(d,L)=\frac{E[\min(L,Z)]-E[\min(d,Z)]}{L-d}. \tag{16}$$

with

$$E[\min(x,Z)]=x-E[(x-Z)_+], \tag{17}$$

which follows from the identity $(Z-x)_+=x-E[(x-Z)_+]$. The remaining expected value is evaluated using (12) and (14) according to the integral

$$E[(x-Z)_+]=\int_0^x C\left[F_X(\frac{\lambda}{\lambda_X}t),F_Y(\frac{\lambda}{\lambda_Y}[x-t])\right]dt. \tag{18}$$

6. STATISTICAL ESTIMATION OF BEV COPULA

Instead of parametric models, it is also possible to use non-parametric inference for the dependence function $A(t)$. Our statistical analysis is based on the estimator by Capéeraà, Fougères and Genest [36]. Recall the simplified definition of this estimator introduced in Segers [37]. Let $(X_i,Y_i), i=1,...,n$, be independent random pairs with standard exponential margins and the same joint survivor function given by

$$S(x,y)=P(X_i>x,Y_i>y)=\exp\left[-(x+y)A\left(\frac{y}{x+y}\right)\right],\quad i=1,...,n. \tag{19}$$

Set $\xi_i(t)=\min\left\{\frac{X_i}{1-t},\frac{Y_i}{t}\right\}, 0\le t\le 1,$ with $\xi_i(0)=X_i$ and $\xi_i(1)=Y_i$, $i=1,...,n$. Since the distribution of $\xi(t)=\min\left\{\frac{X}{1-t},\frac{Y}{t}\right\}$ is exponential with mean $E[\xi(t)]=A(t)^{-1}$, one has

$$E[\ln\{\xi(t)\}]=-\ln\{A(t)\}-\gamma,\quad 0\le t\le 1, \tag{20}$$

where $\gamma = 0.57722$ is Euler's constant. Based on (20) a naive estimator for $\ln\{A(t)\}$ is

$$\ln\left\{\hat{A}_n^{CFG}(t)\right\} = -\tfrac{1}{n}\sum_{i=1}^{n}\ln\left\{\xi_i(t)\right\} - \gamma\,. \tag{21}$$

To fulfill the constraints $A(0) = A(1) = 1$, a modification designed to accommodate for these constraints is the CFG estimator by Capéeraà, Fougères and Genest [36] given for arbitrary $p(t)$ with $p(0) = 1,\ p(1) = 0$ by

$$\ln\left\{\hat{A}_n^{CFG}(t,p)\right\} = \ln\left\{\hat{A}_n^{CFG}(t)\right\} - p(t)\cdot\ln\left\{\hat{A}_n^{CFG}(0)\right\} - \left[1 - p(t)\right]\cdot\ln\left\{\hat{A}_n^{CFG}(1)\right\}. \tag{22}$$

In practice, the simple choice $p(t) = 1 - t$ is very close to the optimal choice with minimum asymptotic variance of the estimator at independence. In a recent paper by Genest and Segers [38], the following remarkable property of a rank-based version of the CFG estimator has been derived. At independence, the rank-based CFG estimator is shown to be asymptotically more efficient than all of the known rank-based versions of its competitors.

Given the parametric models of the dependence function considered in Section 2, it is now possible to fit them as closely as possible to the CFG estimator (22) (or better its rank-based version) and use them for actuarial evaluation as illustrated in the next Section.

7. A NUMERICAL ILLUSTRATION

The results of the preceding Sections are now illustrated using a data set from insurance practice. Two lines of a newly founded non-life reinsurance captive of a corporate firm, namely product liability (PL) and general liability (GL), have experienced over a period of 18 months simultaneously the 11 losses reported in Table 1.

Table 1. Bivariate sample of insurance losses

Claim Number	PL Claims	GL Claims
1	790	648
2	1'181	988
3	2'027	3'152
4	4'340	4'876
5	10'823	11'593
6	20'057	25'616
7	21'240	72'175
8	108'199	89'281
9	123'016	118'873
10	449'595	222'943
11	1'284'781	1'319'084

Though only a few observations are available, an actuary faces the task to assess the bivariate risk of these lines of business in order to evaluate various risk quantities, in particular bivariate return periods, actuarial premiums and bivariate rates on line of an excess-of-loss reinsurance. Due to the small sample size of the available data, the most recent statistical theory by Genest and Segers [38] for BEV copula is of limited value. In particular, the available simple test to check whether the data actually follows a BEV copula seems too liberal for small sample sizes and not conclusive, even for samples containing up to 100 data points according to the invited plenary talk of Genest [39]. Whether the data of Table 1 comes actually from a BEV copula remains thus a statistical challenge open for future research. On the other hand, the situation is less hopeless if one assumes that the underlying copula is a BEV copula. Indeed, simulations by Genest and co-workers indicate that the most recent rank-based version of the CFG estimator defined in Genest and Segers [38] remains more efficient than its rank-based competitors for small samples and therefore it should be the preferred choice for estimation.

What about the marginal distributions? It is well-known that in situations involving large losses one should consider marginal distributions with Pareto right tails, which are consistent with Extreme Value Theory (e.g. Embrechts et al. [40]). Good fitting results with similar data sets of larger sample sizes have been obtained so far using a right-tailed Pareto lognormal (PLN) distribution with 3 parameters considered in Reed [41], Reed and Jorgensen [42], and Hürlimann [43] in actuarial science. This distribution is defined by the analytical expression

$$F_X(x;\alpha,\upsilon,\tau)=\Phi\left(\frac{\ln x-\nu}{\tau}\right)-x^{-\alpha}\exp\left(\alpha\nu+\tfrac{1}{2}(\alpha\tau)^2\right)\cdot\Phi\left(\frac{\ln x-\nu}{\tau}-\alpha\tau\right), \quad (23)$$

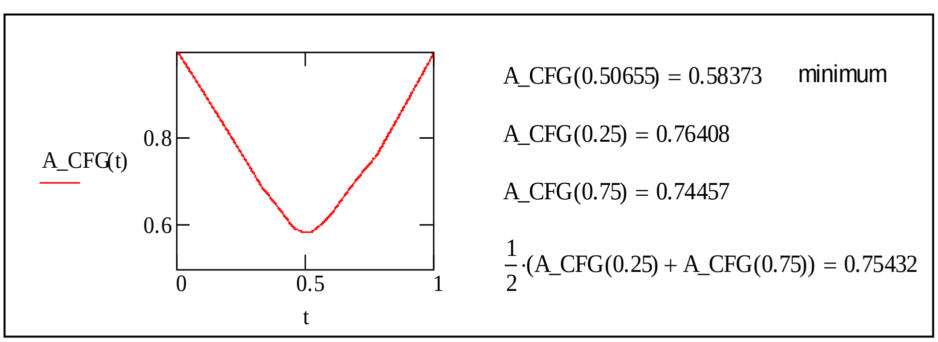
with $\Phi(x)$ the standard normal distribution. Table 2 summarizes the statistical fit of the marginal distributions. The used goodness-of-fit measures are chi-square and Cramér-von Mises like statistics, which are based on the respective differences between the percent rank and the fitted probability distribution evaluated at the level of a given observed claim. The percent rank is defined to be the rank of a value in a data set as a percentage of the data set (e.g. “percentrank” function in Microsoft EXCEL Software).

Table 2. Statistical fit of marginal distributions

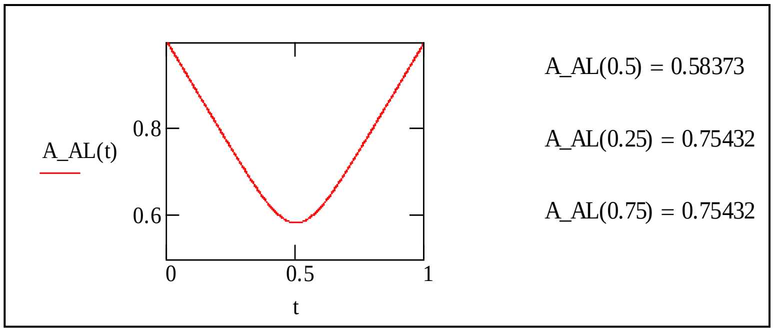
Parameters	PL distribution	GL distribution
α	**0.54995**	**0.70000**
ν	**8.36748**	**8.36387**
τ	**1.66452**	**2.36316**
Goodness-of-fit statistics		
chi-square value χ_2	**1.67**	**0.57**
Cramér - von Mise K	**3.24**	**4.35**

To fit a BEV copula to the data set of Table 1, the above PLN marginal fitted distributions are used and the dependence function $A(t)$ of the BEV copula is fitted using the CFG estimator of Section 6, which is then approximated using as parametric model the Gumbel model and its asymmetric logistic enlargement as defined in Section 2. The graphs of the estimated dependence functions are displayed in the Figure 1. It turns out that the one-parameter Gumbel model and the two-parameter asymmetric logistic model with $\phi=\theta$ both almost coincide with the CFG estimator. Since no significant differences in calculated actuarial figures can be detected in the Tables 3 to 5 below, the parsimonious Gumbel model can be recommended in our small sample size situation.

CGF estimator:



Asymmetric logistic model: $r=4.794, \quad \varphi=\theta=0.98591$



Gumbel model: $r = 4.47676$

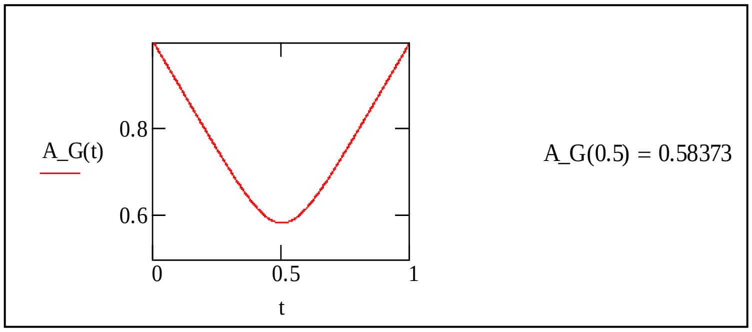


Figure 1. Graphs of the fitted dependence functions

The impact of the chosen fitted models on the bivariate return periods of Section 3 is exhibited in Table 3. Since there are 11 losses over 18 months, a straightforward estimate of the frequency parameter in Proposition 3.1 is $f = 11/5 = 7.33$. The actuarial valuation of various retained XL reinsurance layers following Section 4 is shown in Table 4. Finally, Table 5 lists bivariate ROL for XL covers calculated according to Section 5.

Table 3. Return periods for BEV copulas

Asymmetric logistic model:

Event (X>x OR Y>y) **x = PL claim**	**y = GL claim** **5'000'000**	**10'000'000**	**15'000'000**	**20'000'000**	**25'000'000**
5'000'000	3.98	4.28	4.31	4.32	4.33
10'000'000	4.74	5.96	6.21	6.28	6.31
15'000'000	4.92	6.88	7.54	7.75	7.83
20'000'000	4.98	7.36	8.46	8.89	9.07
25'000'000	5.01	7.61	9.09	9.78	10.10

Event (X>x AND Y>y) **x = PL claim**	**y = GL claim** **5'000'000**	**10'000'000**	**15'000'000**	**20'000'000**	**25'000'000**
5'000000	5.69	8.35	10.87	13.20	15.38
10'000000	6.99	8.85	11.11	13.34	15.47
15'000000	8.36	9.61	11.52	13.58	15.62
20'000'000	9.64	10.52	12.06	13.92	15.85
25'000'000	10.82	11.46	12.71	14.34	16.14

Gumbel model:

Event (X>x OR Y>y) **x = PL claim**	**y = GL claim** **5‘000‘000**	**10‘000‘000**	**15‘000‘000**	**20‘000‘000**	**25‘000‘000**
5‘000‘000	3.98	4.29	4.33	4.34	4.35
10‘000‘000	4.74	5.97	6.24	6.31	6.34
15‘000‘000	4.94	6.88	7.55	7.77	7.86
20‘000‘000	5.00	7.36	8.46	8.91	9.10
25‘000‘000	5.03	7.62	9.09	9.79	10.13

Event (X>x AND Y>y) **x = PL claim**	**y = GL claim** **5‘000‘000**	**10‘000‘000**	**15‘000‘000**	**20‘000‘000**	**25‘000‘000**
5‘000‘000	5.69	8.28	10.75	13.04	15.18
10‘000‘000	6.98	8.83	11.04	13.22	15.30
15‘000‘000	8.31	9.61	11.48	13.50	15.49
20‘000‘000	9.56	10.51	12.05	13.87	15.75
25‘000‘000	10.72	11.43	12.71	14.32	16.07

Table 4. Actuarial valuation of retained XL reinsurance layers

	severity		**aggregate claims**		**actuarial**
model	**mean**	**st. deviation**	**mean**	**st. deviation**	**premium**
25M xs 0M					
asymmetric logistic	983'698	3'794'487	14'427'572	15'012'160	21'933'652
Gumbel	981'367	3'790'298	14'393'390	14'994'391	21'890'586
50M xs 0M					
asymmetric logistic	1'329'639	6'241'469	19'501'374	24'439'384	31'721'066
Gumbel	1'326'535	6'234'720	19'455'849	24'411'629	31'661'663
100M xs 0M					
asymmetric logistic	1'787'978	10'234'362	26'223'683	39'788'263	46'117'815
Gumbel	1'783'872	10'223'267	26'163'461	39'743'700	46'035'311
200M xs 0M					
asymmetric logistic	2'396'279	16'745'519	35'145'424	64'783'743	67'537'296
Gumbel	2'390'873	16'727'845	35'066'141	64'713'810	67'423'046
500M xs 0M					
asymmetric logistic	3'518'714	32'138'310	51'607'812	123'815'859	113'515'742
Gumbel	3'510'980	32'105'167	51'494'373	123'686'463	113'337'604

Table 5. Bivariate ROL for XL reinsurance layers for upper limit of $L = 100$ *Mio.*

Priority	**Cover**	**Bivariate ROL Asymmetric Logistic**	**Bivariate ROL Gumbel**
1'000'000	99'000'000	1.058%	1.056%
2'000'000	98'000'000	1.006%	1.004%
3'000'000	97'000'000	0.970%	0.968%
4'000'000	96'000'000	0.941%	0.939%
5'000'000	95'000'000	0.918%	0.916%
7'500'000	92'500'000	0.871%	0.870%
10'000'000	90'000'000	0.836%	0.834%
20'000'000	80'000'000	0.743%	0.741%

8. REFERENCES

[1] A.C. Cebrian, M. Denuit and P. Lambert, Analysis of bivariate tail dependence using extreme value copulas: An application to the SOA medical large claims database, *Belgian Actuarial Bulletin* **3(1)**, 33-41 (2003).

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

[3] W. Hürlimann, Fitting return periods for largest claims with a Fréchet copula: a case study, In: 28th International Congress of Actuaries, Paris (2006), URL: http://www.ica2006.com/Papiers/2004/2004.pdf.

[4] J. Pickands, Multivariate extreme value distributions, *Bulletin of the International Statistical Institute, Proceedings of the 43rd Session*, Buenos Aires, 859-878 (1981).

[5] J.A. Tawn, Bivariate extreme value theory: models and estimation, *Biometrika* **75**, 397-415 (1988).

[6] E.L. Lehmann, Some concepts of dependence, *The Annals of Mathematical Statistics* **37**, 1137-1153 (1966).

[7] A.I. Garralda Guillem, Structure de dépendence des lois de valeurs extrêmes bivariées, *Comptes Rendus de l'Académie des Sciences de Paris* **330**, 593-596 (2000).

[8] T.P. Hutchinson and C.D. Lai, Continuous bivariate distributions, emphasizing applications, Rumsby Scientific, Adelaide (1990).

[9] W. Hürlimann, Hutchinson-Lai's conjecture for bivariate extreme value copulas, *Statistics and Probability Letters* **61(2)**, 191-198 (2003).

[10] W. Hürlimann, Properties and measures of dependence for the Archimax copula, *Advances and Applications in Statistics* 5(2), 125-143 (2005).

[11] S. Kotz and S. Nadarajah, Extreme Value Distributions, Imperial College Press, London (2000).

[12] J. Beirlant, Y. Goegebeur, J. Segers and J. Teugels, Statistics of Extremes: Theory and Applications. J. Wiley, New York (2004).

[13] E.J. Gumbel, Bivariate exponential distributions, *J. American Statistical Association* **55**, 698-707 (1960).

[14] C. Genest and L.P. Rivest, A characterization of Gumbel's family of extreme value distributions, *Statistics and Probability Letters* **8**, 207-211 (1989).

[15] G. Salvadori, Bivariate return periods via 2-Copulas, Statistical Methodology 1, 129-144 (2004).

[16] G. Salvadori, C. DeMichele, N.T. Kottegoda and R. Rosso, Extremes in nature: an approach using copulas, WST Library Series 56, Springer (2007).

[17] J.T. Shiau, Return period of bivariate distributed extreme hydrological events, *Stochastic Environmental Research and Risk Assessment* **17(1-2)**, 42-57 (2003).

[18] J. Galambos, The asymptotic theory of extreme order statistics, J. Wiley, New York (1978).

[19] C. Genest and L.-P. Rivest, Statistical inference procedures for bivariate Archimedean Copulas, *J. Amer. Statist. Assoc.* **88(423)**, 1034-1043 (1993).

[20] C. Genest and L.-P. Rivest, On the multivariate probability integral transform, *Statist. Probab. Lett.* **53**, 391-399 (2001).

[21] P. Capéraà, P.-L. Fougères and C. Genest, A stochastic ordering based on a decomposition of Kendall's tau, In: Benes, V. and J. Stepân (Eds.), Distributions with given marginals and moment problems, 81-86. Kluwer Acad. Publ., Dordrecht (1997).

[22] K. Ghoudi, K. Khoudraji and L.-P. Rivest, Propriétés statistiques de valeurs extrêmes bidimensionnelles, *Canad. J. Statist.* **26**, 187-197 (1998).

[23] A. Chakak and M. Ezzerg, Bivariate contours of copulas, *Comm. Statist. Simulation Comput.* **29(1)**, 175-185 (2000).

[24] A. Chakak and L. Imlahi, Multivariate probability integral transformation: application to maximum likelihood estimation, *Rev. R. Acad. Cien. Serie A. Mat.* **95(2)**, 201-212 (2001).

[25] R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodriguez-Lallena and M. Ùbeda-Flores, Kendall distribution functions, *Statist. Probab. Lett.* **65**, 263-268 (2003).

[26] C. Genest and J.-C. Boies, Detecting dependence with Kendall plots, *Amer. Statist.* **57**, 275-284 (2003).

[27] C. Genest, J.-F. Quessy and B. Rémillard, Tests of serial independence based on Kendall's process, *Canad. J. Statist.* **30(3)**, 1-21 (2002).

[28] C. Genest, J.-F. Quessy and B. Rémillard, Goodness-of-fit procedures for copula models based on the probability integral transformation, *Scand. J. Statist.* **33**, 337-366 (2006).

[29] N. Kolev, U. Ulisses dos Anjos and B.V. de M. Mendes, Copulas: a review and recent developments, *Stochastic Models* **22(4)**, 617-660 (2006).

[30] F. Belzunce, A. Castaño, A. Olvera-Cervantes and A. Suàrez-Llorens, Quantile curves and dependence structure for bivariate distributions, *Comput. Statist. Data Anal.* **51(10)**, 5112-5129 (2007).

[31] E.C. Brechmann, Hierarchical Kendall copulas: properties and inference, Preprint (2013), URL: http://arxiv.org/pdf/1202.1998v4.pdf

[32] W.-R. Heilmann, On the impact of the independence of risks on stop-loss premiums, *Insurance: Mathematics and Economics* **5**, 197-199 (1986).

[33] W. Hürlimann, Bivariate distributions with diatomic conditionals and stop-loss transform of random sums. *Statistics and Probability Letters* **17**, 329-35 (1993).

[34] W. Hürlimann, On best stop-loss bounds for bivariate sums by known marginal means, variances and correlation, *Bulletin Swiss Association of Actuaries*, 111-34 (1998).

[35] J. Dhaene and M.J. Goovaerts, Dependency of risks and stop-loss premiums, *ASTIN Bulletin* **26**, 201-212 (1996).

[36] P. Capéeraà, A.L. Fougères and C. Genest, A nonparametric estimation procedure for bivariate extreme value copulas, Biometrika 84, 567-577 (1997).

[37] J. Segers, Non-parametric inference for bivariate extreme-value copulas, In: M. Ahsanullah and S. Kirmani (Eds.), Extreme Value Distributions, Nova Science Publishers, pp. 181–203 (2007).

[38] C. Genest and J. Segers, Rank-based inference for bivariate extreme-value copulas, *Annals of Statistics* **37(5B)**, 2597-3097 (2009).

[39] C. Genest, Accounting for Extreme-Value Dependence in Multivariate Data. Invited plenary talk, 38th ASTIN Colloquium, Manchester (2008). URL: http://www.actuaries.org/ASTIN/Colloquia/Manchester/Programme_EN.cfm

[40] P. Embrechts, C. Klüppelberg and Th. Mikosch, Modelling Extremal Events for Insurance and Finance. Springer-Verlag, Berlin (1997).

[41] W.J. Reed, The Pareto law of incomes – an explanation and an extension, Physica A319, 469-486 (2003).

[42] W.J. Reed and M. Jorgensen, The double Pareto-lognormal distribution – A new parametric model for size distribution, Communications in Statistics – Theory & Methods 33(8), 1733-1753 (2004).

[43] W. Hürlimann, Pareto type distributions and excess-of-loss reinsurance, Int. J. Research and Reviews in Applied Sciences 18(3), (2014).