SORT 39 (2) July-December 2015, 209-230

On the bivariate Sarmanov distribution and copula. An application on insurance data using truncated marginal distributions

Zuhair Bahraoui[1], Catalina Bolancé[1], Elena Pelican[2] and Raluca Vernic[2]

Abstract

The Sarmanov family of distributions can provide a good model for bivariate random variables and it is used to model dependency in a multivariate setting with given marginals. In this paper, we focus our attention on the bivariate Sarmanov distribution and copula with different truncated extreme value marginal distributions. We compare a global estimation method based on maximizing the full log-likelihood function with the estimation based on maximizing the pseudo-log-likelihood function for copula (or partial estimation). Our aim is to estimate two statistics that can be used to evaluate the risk of the sum exceeding a given value. Numerical results using a real data set from the motor insurance sector are presented.

MSC: MSC62-07 MSC62E15 and MSC62F10

Keywords: Bivariate Sarmanov distribution, truncated marginal distributions, copula representation, risk measures.

1. Introduction

Modelling bivariate variables that represent economic losses is not a straightforward task. To analyse such data, the usual approach involves fitting univariate distributions to both marginals and then considering the corresponding theoretical bivariate distribution for the entire data set. However, this procedure might not be successful if the marginals present different distribution types or if the dependency structure of the theoretical bivariate distribution is inappropriate for the real data. Furthermore, given the shape of the likelihood function or moments, estimating the parameters can be challenging.

[1] Department of Econometrics, Riskcenter-IREA, University of Barcelona, Av. Diagonal, 690, 08034 Barcelona, Spain.

[2] Faculty of Mathematics and Computer Science, Ovidius University of Constanta, 124 Mamaia Blvd, 900527 Constanta, Romania.
Received: December 2014
Accepted: April 2015

On the other hand, when the marginals follow extreme value distributions, in some cases we have infinite moments. In an economic context, this means that the loss amount is unlimited. However, in practice, this is an unrealistic scenario.

In this paper, we limit ourselves to the Sarmanov family of distributions, originally introduced in its bivariate form by Sarmanov (1966) to join given marginals. This distribution has also been proposed in a more general form in the field of physics by cohen (1984), its multivariate version was suggested by Lee (1996) and generalised further by Bairamov et al. (2001) and Bairamov et al. (2011). Recently, the Sarmanov distribution has attracted interest in other fields (see, for example, Danaher, 2007; Gómez-Déniz and Pérez-Rodríguez, 2015), including finance and insurance. Thus, Hernández-Bastida et al. (2009) and Hernández-Bastida and Fernández-Sánchez (2013) used the bivariate Sarmanov distribution for evaluating premiums in insurance compound models, while further applications related to the theory of ruin were presented by Yang and Hashorva (2013). Furthermore, Hashorva and Ratovomirija (2015) have analysed the Sarmanov distribution with mixed Erlang marginal distributions and have used it for capital allocation. In general, this family of distributions is useful for analysing multivariate loss data, whose marginal distributions may be of the extreme value type or may present very different behaviours. We propose a global estimation (GE) method for the parameters of the Sarmanov distribution with right truncated extreme value marginal distributions.

The bivariate Sarmanov copula is derived from the bivariate Sarmanov distribution and can be a good, quite simple alternative for representing dependency. A copula is a function that relates a bivariate distribution function to its univariate marginal distribution functions, thus allowing the structure of dependence between variables to be fitted separately from the marginal distributions. Specifically, we focus our attention on the bivariate Sarmanov distribution and copula with different log-types of truncated marginal distributions: truncated log-normal, mixture of truncated log-normals and truncated log-logistic. The proposed models may be useful for measuring the risk of loss.

When analising data that represent univariate losses, the univariate distribution that generates the observations is often an extreme value distribution and, therefore, the mean or variance (first or second moment) of the corresponding random variable can be infinite. In finance and insurance, for quantifying the risk it is useful to assume a finite value for the first two moments of the distribution, leading to the right truncation of the distribution of the random variable analised, which was the procedure adopted in this paper. Furthermore, we use a bivariate Sarmanov distribution that requires marginal distributions with finite first moment.

Using a real data set from the motor insurance sector, we compare the estimated risk of loss evaluated for the bivariate Sarmanov distribution with truncated extreme value marginal distributions whose parameters result by the GE method, with the estimation of the same risk obtained after Monte Carlo simulation from the corresponding copula (as examples of fitting alternative copulas and marginals on this data set see, Bolancé et al., 2014; Bahraoui et al., 2014).

The paper is structured as follows: in Section 2 we present two truncated log-normal-type univariate distributions, plus the heavier-tailed truncated log-logistic (Champernowne) distribution, for which we also obtained the first and second moments. In Section 3 we introduce the bivariate Sarmanov distribution and its copula representation, and discuss the parameters estimation. Some comments on the evaluation of two statistics that are used to quantify the risk of loss (Value at Risk - VaR and Tail Value at Risk - TVaR) are presented in Section 4. Finally, in Section 5 we present the results of the proposed fits and risk estimations. Section 6 concludes.

2. Some univariate truncated distributions

We begin by introducing some notations and some univariate truncated distributions to be used as marginals for the bivariate Sarmanov distribution and copula in Section 3.

Let $\mathbf{X} = (X_1, X_2)$ be a bivariate random vector that represents two dependent losses. The random variable (r.v.) $S = X_1 + X_2$ is the total loss and we are interested in measuring the risk associated with the distribution of S; for this, we need to consider both the joint distribution of $\mathbf{X}$ and the marginal distributions of X_1, X_2.

In this section, we analise the probability distribution function (pdf), the cumulative distribution function (cdf) and the first two moments of three distributions that can be useful to model losses: the truncated log-normal, the mixture of two truncated log-normals and the truncated log-logistic, also known as the Champernowne distribution; we let m and M be the truncation points[1] on the left and right side, respectively.

2.1. Truncated log-normal distributions

Let $\varphi(\cdot)$ and $\Phi(\cdot)$ denote the pdf and the cdf, respectively, of the standard normal $N(0,1)$ distribution. To denote the pdf of the general normal $N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma > 0$ distribution, we use the same symbol φ emphasizing the parameters, i.e. $\varphi(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$. The truncated normal distribution $TN(\mu, \sigma^2; m, M), \mu \in \mathbb{R}, \sigma > 0$, with truncation points $m < M$, has the pdf

$$f_{TN}(x) = \frac{\varphi(x; \mu, \sigma^2)}{\Phi(A) - \Phi(a)} = \frac{1}{(\Phi(A) - \Phi(a))\,\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, m \leq x \leq M,$$

where $A = \frac{M-\mu}{\sigma}, a = \frac{m-\mu}{\sigma}$. Its expected value and variance are given respectively, by (see, for example, Kotz et al.,2000)

1. In our numerical application we assume $m = 0$.

$$\xi = \mu + \frac{\varphi(a) - \varphi(A)}{\Phi(A) - \Phi(a)}\sigma,\ \delta^2 = \left(1 + \frac{a\varphi(a) - A\varphi(A)}{\Phi(A) - \Phi(a)} - \left(\frac{\varphi(a) - \varphi(A)}{\Phi(A) - \Phi(a)}\right)^2\right)\sigma^2.$$

We recall that a random variable (r.v.) X follows a log-normal distribution $LN\left(\mu, \sigma^2\right)$ if $\ln X \sim N\left(\mu, \sigma^2\right)$, having hence the pdf $f_X(x) = \frac{1}{x}\varphi\left(\ln x; \mu, \sigma^2\right)$ and cdf $f_X(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right), x > 0$. Moreover, we say that X follows a truncated log-normal distribution $TLN\left(\mu, \sigma^2; m, M\right)$ with truncation points $0 < m < M$, if $\ln X \sim TN\left(\mu, \sigma^2; \ln m, \ln M\right)$; hence, its pdf is $f_X(x) = \frac{1}{x}\frac{\varphi\left(\ln x; \mu, \sigma^2\right)}{\Phi(B) - \Phi(b)}$, where $B = \frac{\ln M - \mu}{\sigma}, b = \frac{\ln m - \mu}{\sigma}$.

Proposition 1 *If $X \sim TLN\left(\mu, \sigma^2; m, M\right), 0 < m < M$, its first two moments are given by*

$$E[X] = e^{\mu + \frac{\sigma^2}{2}}\frac{\Phi(C) - \Phi(c)}{\Phi(B) - \Phi(b)},\ E\left[X^2\right] = e^{2\left(\mu + \sigma^2\right)}\frac{\Phi(D) - \Phi(d)}{\Phi(B) - \Phi(b)},$$

where $C = B - \sigma, c = b - \sigma, D = B - 2\sigma, d = b - 2\sigma$.

Proof Changing variable $y = \ln x$, we obtain

$$E[X] = \int_m^M \frac{x}{x}\frac{\varphi\left(\ln x; \mu, \sigma^2\right)}{\Phi(B) - \Phi(b)}dx = \int_{\ln m}^{\ln M}\frac{\varphi\left(y; \mu, \sigma^2\right)e^y}{\Phi(B) - \Phi(b)}dy$$

$$= \frac{e^{\mu + \frac{\sigma^2}{2}}}{\Phi(B) - \Phi(b)}\int_{\ln m}^{\ln M}\varphi\left(y; \mu + \sigma^2, \sigma^2\right)dy,$$

which immediately yields the stated formula of $E[X]$. The formula of $E\left[X^2\right]$ results in a similar way. ∎

2.2. Mixtures of two truncated log-normal distributions

Consider two truncated normal distributions $TN(\mu_i, \sigma_i^2; m, M), \mu_i \in \mathbb{R}, \sigma_i > 0, i = 1, 2$, having the same truncation points $m < M$. Then, denoting their mixture by $TN_{mixt}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, r; m, M\right), r \in (0, 1)$, its pdf has the form

$$f_{TN_{mixt}}(x) = r\frac{\varphi\left(x; \mu_1, \sigma_1^2\right)}{\Phi(A_1) - \Phi(a_1)} + (1 - r)\frac{\varphi\left(x; \mu_2, \sigma_2^2\right)}{\Phi(A_2) - \Phi(a_2)}, m \le x \le M,$$

where $A_i = \frac{M - \mu_i}{\sigma_i}, a_i = \frac{m - \mu_i}{\sigma_i}, i = 1, 2$, and r is the mixing parameter.

Similarly, we say that the r.v. X follows a mixture of two truncated log-normal distributions $TLN_{mixt}\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,r;m,M\right),m>0$, if its pdf is

$$f_{TLN_{mixt}}(x)=r\frac{\varphi\left(\ln x;\mu_1,\sigma_1^2\right)}{x\left(\Phi(B_1)-\Phi(b_1)\right)}+(1-r)\frac{\varphi\left(\ln x;\mu_2,\sigma_2^2\right)}{x\left(\Phi(B_2)-\Phi(b_2)\right)},m\leq x\leq M,$$

with $B_i=\frac{\ln M-\mu_i}{\sigma_i},b_i=\frac{\ln m-\mu_i}{\sigma_i},i=1,2$. In this case, $\ln X\sim TN_{mixt}\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,r;m',M'\right)$, where $m'=\ln m,M'=\ln M$.

To obtain the moments of the above mixtures of truncated distributions, we note that the pdf of such a mixture of distributions is of the form $f(x)=rf_1(x)+(1-r)f_2(x)$, where f_1 and f_2 are themselves pdf's. If we denote by X_i a r.v. having pdf f_i and by X a r.v. with pdf f, then the first two moments of the mixed distribution results as

$$E[X]=rE[X_1]+(1-r)E[X_2],\ E\left[X^2\right]=rE\left[X_1^2\right]+(1-r)E\left[X_2^2\right],$$

from where a straightforward calculation yields the variance

$$Var[X]=rVar[X_1]+(1-r)Var[X_1]+r(1-r)\left(E[X_1]-E[X_2]\right)^2.$$

Using these formulas, the first moments of the TN_{mixt} and TLN_{mixt} distributions are immediate.

Moreover, we also note that fitting a truncated log-normal distribution or a mixture of two truncated log-normal distributions to a data set, is the same as fitting a truncated normal distribution or, correspondingly, a mixture of two truncated normal distributions to the log-data set.

2.3. Champernowne (log-logistic) distribution

Introduced by Champernowne in 1952 (see, Champernowne, 1952), the log-logistic distribution is the distribution of a r.v. whose logarithm follows a logistic distribution. In economics, where it is also known as the Fisk distribution (see, Fisk, 1961), it is used to model the distribution of wealth or income. Its shape is similar to the log-normal distribution, but it has heavier tails; moreover, as an asymptotic behaviour, it converges towards a Pareto distribution in the tail (see, Buch-Larsen et al., 2005). Denoted by $Ch(\alpha,H),\alpha,H>0$, its pdf is defined by

$$f_{Ch}(x)=\frac{\alpha H^{\alpha}x^{\alpha-1}}{\left(x^{\alpha}+H^{\alpha}\right)^2},x\geq 0,$$

having cdf $F_{Ch}(x)=\frac{x^{\alpha}}{x^{\alpha}+H^{\alpha}}, x\geq 0$, expected value $\frac{\pi H}{\alpha}\left(\sin\frac{\pi}{\alpha}\right)^{-1}$, for $\alpha>1$, and variance $\frac{\pi H^2}{\alpha^2}\left(\sin\frac{\pi}{\alpha}\right)^{-1}\left(\left(\cos\frac{\pi}{\alpha}\right)^{-1}-\pi\left(\sin\frac{\pi}{\alpha}\right)^{-1}\right)$, for $\alpha>2$. Note that H is a scale parameter and the median of the distribution, while α is a shape parameter.

We also consider the truncated form $TCh(\alpha,H;M), \alpha,H,M>0$, having pdf[2]

$$f_{TCh}(x)=\alpha\left(M^{\alpha}+H^{\alpha}\right)\left(\frac{H}{M}\right)^{\alpha}\frac{x^{\alpha-1}}{\left(x^{\alpha}+H^{\alpha}\right)^2}, 0\leq x\leq M. \tag{1}$$

Its moments do not have a closed form, but they can be expressed in terms of the hypergeometric function ${}_2F_1$ defined for $|z|<1$ by the following integral or power series

$${}_2F_1(a,b;c;z)=\frac{1}{B(b,c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}dt=\sum_{k=0}^{\infty}\frac{(a)_k(b)_k}{(c)_k}\frac{z^k}{k!},$$

where

$$(q)_k=\begin{cases}1, & k=0\\ q(q+1)\cdots(q+k-1), & k>0\end{cases}, c\notin\{0,-1,-2,...\},$$

and $B(a,b)=\int_0^1 t^{a-1}(1-t)^{b-1}dt$ is the Beta function.

Proposition 2 *Letting $X\sim TCh(\alpha,H;M)$, its first two moments are given by*

$$E[X]=\frac{\alpha M}{\alpha+1}\left(1+\left(\frac{M}{H}\right)^{\alpha}\right){}_2F_1\left(2,1+\frac{1}{\alpha};2+\frac{1}{\alpha};-\left(\frac{M}{H}\right)^{\alpha}\right), \tag{2}$$

$$E\left[X^2\right]=\frac{\alpha M^2}{\alpha+2}\left(1+\left(\frac{M}{H}\right)^{\alpha}\right){}_2F_1\left(2,1+\frac{2}{\alpha};2+\frac{2}{\alpha};-\left(\frac{M}{H}\right)^{\alpha}\right). \tag{3}$$

Proof We evaluate the expected value of X by changing variable $x=My^{1/\alpha}$ in

$$\begin{aligned}E[X]&=\alpha\left(M^{\alpha}+H^{\alpha}\right)\left(\frac{H}{M}\right)^{\alpha}\int_0^M\frac{x^{\alpha}}{\left(x^{\alpha}+H^{\alpha}\right)^2}dx\\&=\alpha\left(M^{\alpha}+H^{\alpha}\right)\left(\frac{H}{M}\right)^{\alpha}\int_0^1\frac{\left(My^{1/\alpha}\right)^{\alpha}}{\left(\left(My^{1/\alpha}\right)^{\alpha}+H^{\alpha}\right)^2}\frac{My^{1/\alpha-1}}{\alpha}dy\end{aligned}$$

2. Since in our application we assume $m=0$, for the sake of simplicity, we only present the properties for $M>0$.

$$= M\left(M^{\alpha}+H^{\alpha}\right)\left(\frac{H}{M}\right)^{\alpha}\left(\frac{M}{H^{2}}\right)^{\alpha}\int_{0}^{1}\frac{y^{1/\alpha+1-1}}{\left(1+\left(\frac{M}{H}\right)^{\alpha}y\right)^{2}}dy$$

$$= M\left(1+\left(\frac{M}{H}\right)^{\alpha}\right)B\left(1+\frac{1}{\alpha},1\right)\ {}_{2}F_{1}\left(2,1+\frac{1}{\alpha};2+\frac{1}{\alpha};-\left(\frac{M}{H}\right)^{\alpha}\right),$$

with the last relation resulting from the definition of the function ${}_2F_1$. Note that

$$B\left(1+\frac{1}{\alpha},1\right)=\frac{\Gamma\left(1+\frac{1}{\alpha}\right)\Gamma\left(1\right)}{\Gamma\left(2+\frac{1}{\alpha}\right)}=\frac{1}{1+\frac{1}{\alpha}}=\frac{\alpha}{1+\alpha},$$

where $\Gamma(a)=\int_0^{\infty}x^{a-1}e^{-x}dx$ denotes the Gamma function. Inserting this result into the last expression of $E[X]$ immediately yields formula (2). Formula (3) results in a similar way. ∎

3. Bivariate Sarmanov distribution

3.1. The general distribution

We say that the random vector $\mathbf{X}=(X_1,X_2)$ follows a bivariate Sarmanov's distribution if its joint pdf is given by (see, Kotz et al., 2000).

$$f_{\mathbf{X}}(x_1,x_2)=f_1(x_1)f_2(x_2)\left(1+\omega\phi_1(x_1)\phi_2(x_2)\right), \tag{4}$$

where $(f_i)_{i=1,2}$ are the corresponding marginal pdf's, $(\phi_i)_{i=1,2}$ are bounded non-constant kernel functions and ω is a real number such that

$$\int_{-\infty}^{\infty}\phi_i(x_i)f_i(x_i)dx_i=0,\ i=1,2,\ \text{and} \tag{5}$$

$$1+\omega\phi_1(x_1)\phi_2(x_2)\geq 0,\ \text{for all}\ (x_1,x_2)\in\mathbb{R}^2. \tag{6}$$

If we denote $\nu_i=\int_{-\infty}^{\infty}x\phi_i(x)f_i(x)dx,\ i=1,2,$ then the covariance and correlation coefficient are given, respectively, by

$$\begin{aligned}\operatorname{cov}(X_1,X_2)&=\omega\,\nu_1\,\nu_2,\\ \operatorname{corr}(X_1,X_2)&=\frac{\omega\,\nu_1\,\nu_2}{\sqrt{Var[X_1]\,Var[X_2]}}.\end{aligned} \tag{7}$$

Note that when ω=0, X_1 and X_2 are independent. As to the choice of the kernel functions ϕ_i, some particular cases satisfying (5) have already been discussed in the literature (see, Lee, 1996), from which we recall:

- $\phi_i = 1-2F_i$, where F_i is the cdf of X_i. In this case, the Sarmanov distribution is known as the Farlie-Gumbel-Morgenstern distribution (see Farlie, 1960), verifying the restrictive condition that the correlation coefficient $\operatorname{corr}(X_1,X_2)$ cannot exceed 1/3 in absolute value. However, in general, the Sarmanov distribution is not restricted by such a condition (see, for example, Shubina and Lee, 2004).
- $\phi_i(x) = e^{-\alpha x} - E\left[e^{-\alpha X_i}\right]$ (we say no more about this form as it did not provide a good fit to our data).
- $\phi_i(x) = x^\alpha - E\left[X_i^\alpha\right]$, assuming that $E\left[X_i^\alpha\right] < \infty$. In this case, $\nu_i = E\left[X_i^{\alpha+1}\right] - E\left[X_i^\alpha\right]E\left[X_i\right]$, if it is finite.

Given its simplicity and better fit for our data, in our study we consider $\phi_i(x) = x^\alpha - E\left[X_i^\alpha\right]$ with $\alpha = 1$, yielding from (7) the correlation

$$\operatorname{corr}(X_1,X_2) = \omega\sqrt{Var\left[X_1\right]Var\left[X_2\right]}. \tag{8}$$

Therefore, assuming that $E\left[X_i\right] < \infty$, in the following we limit ourselves to the pdf form

$$f_{\mathbf{X}}(x_1,x_2) = f_1(x_1)f_2(x_2)\left(1+\omega\left(x_1-E\left[X_1\right]\right)\left(x_2-E\left[X_2\right]\right)\right) \tag{9}$$

that requires the existence of a finite first moment for both marginals. In this case, condition (6) obviously restricts the domain of $f_{\mathbf{X}}$. For simplicity, we preferred to work with truncated marginals, which meant imposing restrictions on the coefficient ω. More precisely, if the support of X_i is $[m_i,M_i], i = 1,2$, then condition (6) yields $l \leq \omega \leq u$, where

$$l = \max\left\{\frac{-1}{\left(M_1-E\left[X_1\right]\right)\left(M_2-E\left[X_2\right]\right)}, \frac{-1}{\left(m_1-E\left[X_1\right]\right)\left(m_2-E\left[X_2\right]\right)}\right\}, \tag{10}$$

$$u = \min\left\{\frac{-1}{\left(M_1-E\left[X_1\right]\right)\left(m_2-E\left[X_2\right]\right)}, \frac{-1}{\left(m_1-E\left[X_1\right]\right)\left(M_2-E\left[X_2\right]\right)}\right\}. \tag{11}$$

Because of the restriction imposed by condition (6), we used marginal distributions with bounded support. Therefore, we considered the truncated distributions presented in Section 2, their choice being driven by the real data to be studied in Section 5.

3.2. Copula representation and simulation

A copula can be defined as a multivariate cdf with standard uniform $[0,1]$ marginals. Then the cdf of a random vector $\mathbf{X} = (X_1, \ldots, X_m)$ can be written in terms of its marginal cdf's using a copula function $C : [0,1]^m \to [0,1]$, as follows $F_{\mathbf{X}}(\mathbf{x}) = C(F_1(x_1), \ldots, F_m(x_m))$; for details on copulas see Nelsen (2006).

Since the Sarmanov bivariate distribution is defined directly from its marginal distributions, its cdf can be immediately expressed as $F_{\mathbf{X}}(\mathbf{x}) = C(F_1(x_1), F_2(x_2))$ using the following copula function

$$C(u_1,u_2) = u_1u_2 + \omega \int_0^{u_1} \phi_1(F_1^{-1}(t))dt \int_0^{u_2} \phi_2(F_2^{-1}(s))ds, 0 \leq u_1,u_2 \leq 1, \tag{12}$$

assuming that F_1^{-1}, F_2^{-1} exist; the corresponding density is

$$c(u_1,u_2) = 1 + \omega\phi_1(F_1^{-1}(u_1))\phi_2(F_2^{-1}(u_2)). \tag{13}$$

Working with the copula representation of the Sarmanov family of distributions has some advantages. The copula representation is straightforward and its estimation procedure is simple. Furthermore, this representation enables us to generate pseudo-random samples from the Sarmanov bivariate distribution. To do this, we first generate values from the Sarmanov copula (12) using the procedure described in Nelsen (2006), which is based on the conditional distribution of a random vector (U_1,U_2) having uniform $[0,1]$ marginals and cdf C, i.e., on $C_{u_1}(u_2) = \Pr(U_2 \leq u_2 | U_1 = u_1)$. Note that

$$C_{u_1}(u_2) = \lim_{\Delta u_1 \to 0^+} \frac{C(u_1 + \Delta u_1, u_2) - C(u_1,u_2)}{\Delta u_1} = \frac{\partial C(u_1,u_2)}{\partial u_1}.$$

The corresponding algorithm is implemented as follows:

1. Generate two independent random values u_1 and z from the uniform $U(0,1)$ distribution.
2. Set $u_2 = C_{u_1}^{(-1)}(z)$, where $C_{u_1}^{(-1)}$ denotes a quasi-inverse of C_{u_1}. The desired pair from the Sarmanov copula is (u_1,u_2).
3. Solving now $F_1(x_1) = u_1$ for x_1 and $F_2(x_2) = u_2$ for x_2 yields the pseudo-random pair (x_1,x_2) simulated from the corresponding bivariate Sarmanov's distribution.

In our case, assuming that the inverses F_1^{-1}, F_2^{-1} exist, the partial derivative of (12) is

$$C_{u_1}(u_2) = u_2 + \omega\phi_1(F_1^{-1}(u_1)) \int_0^{u_2} \phi_2(F_2^{-1}(s))ds.$$

If, in particular, we take the kernel functions $\phi_i(z) = z - E[X_i], i = 1,2$, this gives

$$C_{u_1}(u_2) = u_2 + \omega(F_1^{-1}(u_1) - E[X_1]) \int_0^{u_2} (F_2^{-1}(s) - E[X_2])ds.$$

3.3. Estimation of parameters

Let Θ denote the parameters set of the Sarmanov distribution. First, we estimate the parameters using the maximum likelihood (ML) method, that we named global estimation (GE), based on the random data sample $\{(x_{1j}, x_{2j})\}_{j=1}^n$ consisting of n couples of observations. For estimating the Sarmanov copula, we use the maximum pseudo-likelihood method that we named partial estimation (PE).

3.3.1. Global estimation (GE) method

From density (4), the log-likelihood function to be maximized is

$$\ln L\left(\{(x_{1j}, x_{2j})\}_{j=1}^n; \Theta\right) = \sum_{j=1}^n (\ln f_1(x_{1j}) + \ln f_2(x_{2j}) + \ln(1 + \omega\phi_1(x_{1j})\phi_2(x_{2j}))). \tag{14}$$

The parameters to be estimated are ω, the parameters of f_i, and, eventually, the parameters of ϕ_i. Let θ denote a generic parameter of f_i. The corresponding ML system is

$$\begin{cases} 0 = \dfrac{\partial \ln L}{\partial \theta} = \displaystyle\sum_{j=1}^n \left(\frac{\partial \ln f_1(x_{1j})}{\partial \theta} + \frac{\partial \ln f_2(x_{2j})}{\partial \theta}\right) + \omega \sum_{j=1}^n \frac{1}{1 + \omega\phi_1(x_{1j})\phi_2(x_{2j})} \\ \qquad \times \left(\phi_1(x_{1j}) \dfrac{\partial \phi_2(x_{2j})}{\partial \theta} + \phi_2(x_{2j}) \dfrac{\partial \phi_1(x_{1j})}{\partial \theta}\right), \quad \theta \in \Theta \\ 0 = \dfrac{\partial \ln L}{\partial \omega} = \displaystyle\sum_{j=1}^n \frac{\phi_1(x_{1j})\phi_2(x_{2j})}{1 + \omega\phi_1(x_{1j})\phi_2(x_{2j})}. \end{cases} \tag{15}$$

This system can become quite complex and, therefore, it must be solved using numerical methods that require starting values for the unknown parameters. Such starting values readily result from the method of moments (MM); for example, a value for ω can be obtained from the empirical correlation coefficient, ρ. For more details on this procedure see Pelican and Vernic (2013).

Alternatively, instead of solving the ML system, numerical methods can be used to find the maximum of the log-likelihood function directly. Such an optimization problem can be solved using, for example, a variable neighborhood search (VNS) algorithm (see, Mladenovic and Hansen, 1997).

3.3.2. Partial estimation (PE) method

As discussed above, the GE method can result in cumbersome calculations. For this reason, we suggest comparing it with the alternative method based on maximizing the pseudo-log-likelihood corresponding to the copula representation of the Sarmanov distribution (see, for example, Joe, 1997):

- Using the ML method, we estimate the parameters of the univariate marginal distributions of X_1 and X_2, starting from the corresponding data samples $(x_{1j})_{j=1}^n$ and $(x_{2j})_{j=1}^n$, respectively.
- To obtain the parameter ω of the copula, we use again the ML method on (14), after setting the marginal parameters at the values obtained in the previous step. Note, that it is enough to maximize only the last part of (14), i.e., $\sum_{j=1}^n \ln(1+\omega\phi_1(x_{1j})\phi_2(x_{2j}))$, since the rest does not depend on ω; in fact, this is reduced to applying the ML method to the copula density (13).

4. Evaluating the total risk of loss

Evaluating risk measures for aggregate losses is a challenging task. Let S denote an insurance risk, that is, a non-negative random variable whose cdf is denoted by F_S. A risk measure is generally formulated as a functional from the space of insurance risks to $[0,\infty]$, and its purpose is to provide a single value for the degree of risk associated with the corresponding risk. Among the common risk measures, the Value-at-Risk (VaR) is probably the most frequently adopted. To define it, let $q \in (0,1)$ denote the confidence level required by regulations; then

$$VaR_q[S] := \inf\{x : F_S(x) \geq q\}.$$

The Solvency II Accord drawn up by the EU Commission sets $q = 0.995$ over a one year time horizon.

When heavy tails occur in risk management (see recent episodes of financial instability), a risk measure providing information above a given threshold is recommended. In this respect, the Tail Value-at-Risk (TVaR, also known as the expected shortfall or conditional tail expectation) measure is defined, for $q \in (0,1)$, as

$$TVaR_q[S] := E[S|S > VaR_q[S]].$$

TVaR is considered a coherent risk measure, see Artzner et al. (1999). In some countries, TVaR has already replaced VaR in the regulatory requirements; the current practice is $q = 0.99$ over a one year time horizon.

Let now $S = X_1 + X_2$ be the sum of two possibly dependent insurance risks X_1 and X_2. In this section, our goal is to show how to calculate VaR and TVaR for the risk S when $\mathbf{X} = (X_1, X_2)$ follows the bivariate Sarmanov distribution. Vernic (2014) has analised a closed form for the TVaR of the sum of random variables Sarmanov distributed with exponential marginals. We approach this task in two ways: by direct evaluation and by simulation based on the Sarmanov copula.

4.1. Direct evaluation

To obtain VaR, we must evaluate the cdf of S and then invert it. Letting f_S denote the pdf of S, its cdf results from

$$\begin{aligned} F_S(s) = \int_0^s f_S(x)\,dx &= \int_0^s \int_0^x f_{\mathbf{X}}(x-y, y)\,dydx \\ &= \int_0^s \int_0^1 x f_{\mathbf{X}}(x(1-t), xt)\,dtdx. \end{aligned}$$

Similarly, for TVaR we need

$$\begin{aligned} E[S|S > s_q] &= \frac{1}{1 - F_S(s_q)} \int_{s_q}^{\infty} x f_S(x)\,dx \\ &= \frac{1}{1 - F_S(s_q)} \int_{s_q}^{\infty} \int_0^1 x^2 f_{\mathbf{X}}(x(1-t), xt)\,dtdx, \end{aligned}$$

where $s_q = VaR_q[S]$. As there are no closed formulas for these integrals, they have to be calculated using mathematical software. To do this, we wrote Matlab procedures based on Simpson's composite rule for double integrals (see, for example, Bourden and Faires, 2001), paying special attention to the integrals limits since the marginals are truncated.

4.2. Simulation of the Sarmanov copula

Using the Monte Carlo method, the procedure is as follows:

1. We apply the PE method to the data sample $\{(x_{1j}, x_{2j})\}_{j=1}^{n}$ from which we obtain the estimations of the marginals cdf's, denoted $\hat{F}_i, i = 1, 2$, and the estimated parameter of the Sarmanov copula, $\hat{\omega}$.

2. Using the algorithm described in Section 3.2, we generate the pseudo-random sample $\{(\hat{x}_{1j},\hat{x}_{2j})\}_{j=1}^{r}$ from the bivariate Sarmanov distribution with marginals $\hat{F}_1$ and $\hat{F}_2$, where the sample volume r is large (we used $r = 10000$).
3. We calculate $\hat{s}_j = \hat{x}_{1j} + \hat{x}_{2j}, j = 1,\ldots,r$, and we estimate $VaR_q[S]$ and $TVaR_q[S]$ empirically from the generated pseudo-sample $(\hat{s}_j)_{j=1}^{r}$.

5. Numerical study

We used the bivariate Sarmanov distribution and copula to model a random sample of motor insurance claims consisting of the costs of property damage and medical expenses, kindly provided by a major insurer in Spain for the year 2000. Since the data were collected two years later, in 2002, all the claims included in our sample had been settled. The sample size is $n = 518$ and for each claim, X_1 represents the cost of property damage (including third-part liability), while X_2 represents the cost of medical expenses (i.e., treatments and hospitalization as a result of the accident).

Previously, several bivariate distributions were fitted to these data, the best global fit being provided by the bivariate log-skew-normal distribution with a log-likelihood value of -7323.50 and AIC $= 14663.00$ (see, Bolancé et al., 2008). In an attempt to find a better model, in the numerical part of this paper we fitted the bivariate Sarmanov distribution with different normal-type marginals to the bivariate log-data set. Note that if we fit a bivariate Sarmanov distribution with pdf $f_{\mathbf{Y}}$ to the log-data, then the distribution corresponding to the original data is the bivariate log-Sarmanov with pdf

$$f_{\mathbf{X}}(x_1,x_2) = \frac{1}{x_1 x_2} f_{\mathbf{Y}}(\ln x_1, \ln x_2), x_1,x_2 > 0.$$

This implies that the marginal distributions of the original data are the log-distributions of the corresponding marginals of $\mathbf{Y}$ (in our case, they become of log-normal and log-logistic types).

In the first attempt, we assumed that $\ln X_1$ follows a truncated normal (TN) distribution and we varied the distribution of $\ln X_2$, but since the best fit was provided by the mixture of two truncated normal distributions TN_{mixt} for $\ln X_2$, we decided not to provide details of the other distributions and we concentrated only on the best fit. This choice was also motivated by the fact that when studying separately the marginal distributions of our data set, we noticed that the normal distribution provided a good fit for $\ln X_1$, but unfortunately, this was not the case with $\ln X_2$, which has a less regular histogram; hence, we made use of the property of the Sarmanov distribution of joining different marginals. Alternatively, we also fitted the bivariate Sarmanov distribution with the heavier-tailed Champernowne marginal distributions to the original data.

In Table 1 we show the descriptive statistics for the original data and for the log-data.

Table 1: *Descriptive statistics.*

	Mean	Std.Dev.	Kurtosis	Skewness	Min	Max	Median
	Original data						
X_1	1827.60	6867.81	297.10	15.65	13.00	137936.00	677.00
X_2	283.92	863.17	82.02	8.04	1.00	11855.00	88.00
	Correlation between X_1 and X_2 is 0.73						
	Log-data						
$\ln X_1$	6.44	1.33	0.57	0.21	2.56	11.83	6.52
$\ln X_2$	4.38	1.52	0.45	0.12	0.00	9.38	4.48
	Correlation between $\ln X_1$ and $\ln X_2$ is 0.59						

Table 2: *AIC obtained for different estimated models and methods.*

Method	Marginals	max	max$\times$10	max$\times$100
GE	$X_1 \sim TLN, X_2 \sim TLN_{mixt}$	14839.58	14863.04	14869.42
	$X_1 \sim TCh, X_2 \sim TCh$	14849.26	14878.42	14883.74
PE	$X_1 \sim TLN, X_2 \sim TLN_{mixt}$	14854.79	14868.25	14873.40
	$X_1 \sim TCh, X_2 \sim TCh$	14880.99	14884.31	14884.52

Since we decided to work with truncated distributions (as discussed in Section 3.1), a key issue was the choice of the upper truncation limits, the lower ones being fixed at $m_1 = m_2 = 0$. We started by taking the upper limits as being equal to the maximum observed values, i.e., $M_i = \max_{j=1,\ldots,n} x_{ij}, i = 1,2$. However, this choice most probably underestimates the real risk since it implies the assumption that the probability of a loss greater than the maximum observed is zero, which is not true in practice. Hence, we assumed that the upper truncation limits increase progressively, being equal to 10, 100 and 1000 times the maximum observed values (denoted in the following by max$\times$10, max$\times$100 and max$\times$1000, respectively). We found the results for the truncation limits of max$\times$100 and max$\times$1000 to be similar, hence, we present here only the former, i.e., max$\times$100, which is equivalent to almost eliminating the effect of truncation.

To estimate the parameters using the methods described in Section 3.3, we took the main empirical characteristics as starting values. Then, to compare the different fits, we calculated the corresponding log-likelihood and the Akaike information criterion (AIC) values. AIC is defined by $\text{AIC} = 2(s - \ln L)$, where s is the number of estimated parameters and L is the likelihood function. This criterion penalizes an increased number of parameters, so that the preferred model is the one with the lowest AIC value. In Table 2 we show the AIC obtained for each estimation, while the estimated parameters and their standard errors are shown in Tables 3-6 in the Appendix. It seems that GE yields a slightly better fit than PE, although we observe that the difference between the AICs for GE and PE is small. This is expected since the GE method maximizes the full likelihood, while the PE method maximizes separately the partial likelihoods corresponding to the copula and the marginal distributions. Considering both methods (GE and PE) for all upper truncation limits, it results that the best model is the Sarmanov distribution with a truncated log-normal distribution for X_1 and a mixture of two truncated log-normal distributions for X_2.

In Figures 1 and 2 we plot the VaR and TVaR curves as functions of the confidence level q for $q \geq 0.98$, for all the distributions estimated. In Tables 7-10 in Appendix we also displayed the VaR and TVaR values obtained for the same distributions and for some confidence levels q, compared with the empirical values resulting from data. These values and plots clearly show that for $q \geq 0.95$, the Sarmanov distributions with log-normal-type marginals underestimate the empirical values. Although closer to the empirical curve, this is also the case of the Sarmanov distribution with TCh marginals and an upper truncation limit equal to max, while the other two distributions (i.e., max$\times$10 and max$\times$100) overestimate the empirical values. Therefore, from the point of view of the insurer, only these two last distributions would be of interest.

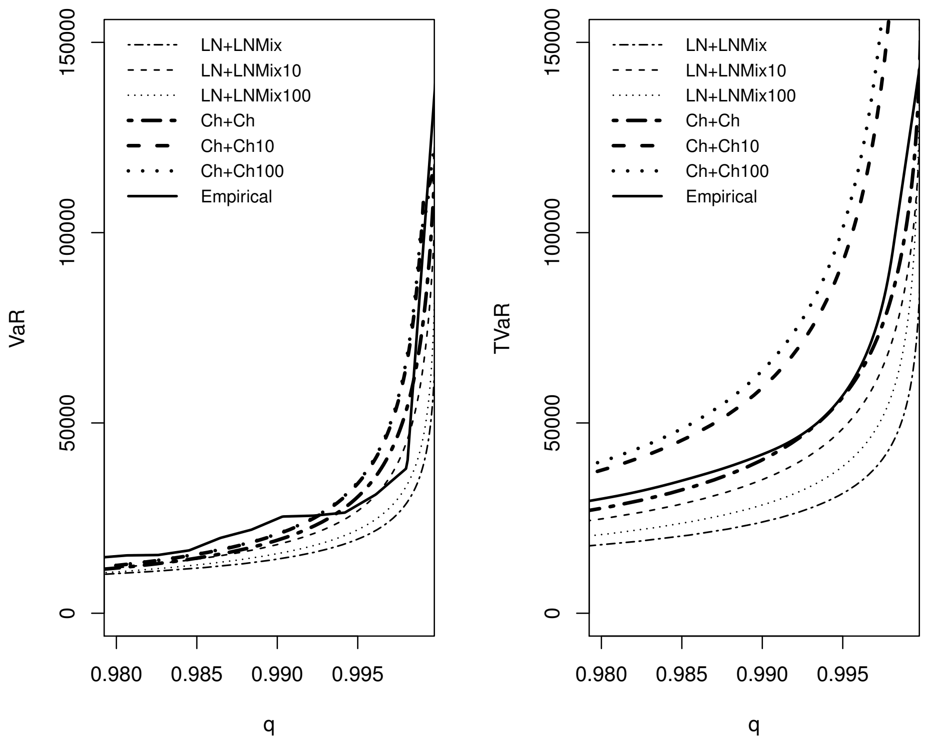

Figure 1: *Estimated VaR and TVaR with GE.*

Note that, the curves resulting from GE and PE methods look similar, although, from Tables 7-10 in Appendix it seems that, in general, PE leads to higher values of VaR and TVaR than those provided by GE.

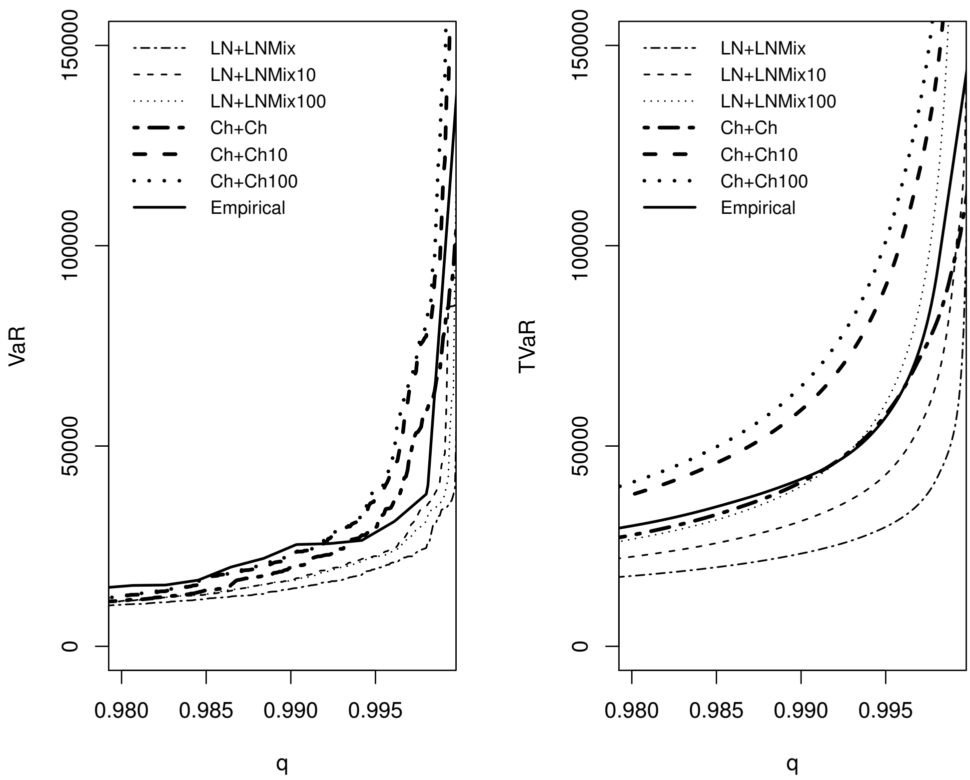

Figure 2: *Estimated VaR and TVaR with PE.*

On the other hand, note that the best globally fitted distribution (in our case, according to AIC, the Sarmanov distributions with LTN and LTN$_{mixt}$ marginals) does not necessarily provide the best model for the risk measures VaR and TVaR, which are defined on the distribution tail – this is also the case with the previously fitted bivariate log-skew-normal distribution, which strongly overestimates the empirical TVaR curve (see Bolancé et al., 2008). For our data set the heavier-tailed Champernowne distribution provides a better model for Sarmanov's marginals when evaluating VaR and TVaR.

6. Conclusions

In this paper, we have proposed the Sarmanov bivariate distribution as a model for bivariate insurance losses and we have illustrated its applicability using a real data set from the motor insurance sector. The choice of this distribution was motivated by its flexible structure that allowed us to join given marginals. From the numerical study, we conclude that the distribution could be a good model for such bivariate insurance data, but special attention should be paid to the choice of the marginal distributions. More specifically, these distributions must fulfill the condition of a real pdf, see (5)-(6), so that truncated marginal distributions can be selected. Moreover, the upper truncation limits have to be carefully fixed so that the real risk values (like VaR or TVaR) should not be underestimated, but also not overestimated to an exaggerated degree.

It should also be noted that a better global fit does not necessarily mean a better fit regarding the evaluation of some tail related risk measures.

As for the choice between GE and PE methods, it seems that GE yields a somewhat better fit than PE, although the differences are very small. However, the application of the GE method might be more time-consuming given the random search involved in the ML solution. Clearly, the complexity of the calculation should be taken into consideration when selecting the most suitable estimation method.

Appendix

Table 3: *GE for* $\ln X_1 \sim TN$ *and* $\ln X_2 \sim TN_{mixt}$ *and different upper truncation limits (standard errors between parentheses).*

	max	max×10	max×100
μ_1	6.4237 (0.0585)	6.4163 (0.0596)	6.4089 (0.0594)
μ_{21}	4.3661 (0.0836)	4.3758 (0.0713)	4.2860 (0.1199)
μ_{22}	4.3771 (0.5458)	4.0157 (0.4906)	4.4288 (0.2702)
σ_1	1.3310 (0.0412)	1.3560 (0.0431)	1.3517 (0.0428)
σ_{21}	1.2420 (0.0833)	1.2938 (0.0569)	1.1653 (0.1140)
σ_{22}	2.9064 (0.8128)	3.0008 (0.3984)	2.0079 (0.2070)
r	0.8079 (0.0889)	0.8456 (0.0383)	0.6733 (0.1348)
ω	0.0404 (0.0210)	0.0214 (0.0188)	0.0162 (0.0180)
$\ln L$	−7411.79	−7423.52	−7426.71
AIC	14839.58	14863.04	14869.42

Table 4: *PE for* $\ln X_1 \sim TN$ *and* $\ln X_2 \sim TN_{mixt}$ *and different upper truncation limits (standard errors between parentheses).*

	max	max×10	max×100
μ_1	6.4439 (0.0587)	6.4437 (0.0553)	6.4437 (0.0587)
μ_{21}	4.3115 (0.1274)	4.1975 (0.0801)	4.2743 (0.1560)
μ_{22}	4.4105 (0.2229)	5.0547 (0.2746)	4.4769 (0.2594)
σ_1	1.3351 (0.0416)	1.3350 (0.0415)	1.3350 (0.0415)
σ_{21}	1.1476 (0.1184)	1.3346 (0.0671)	1.2315 (0.1508)
σ_{22}	1.9488 (0.2144)	1.9550 (0.1994)	2.0372 (0.5330)
r	0.5899 (0.1580)	0.7770 (0.0587)	0.6396 (0.3122)
ω	0.0309 (0.0095)	0.0212 (0.0086)	0.0161 (0.0082)
$\ln L$	−7419.39	−7426.13	−7428.70
AIC	14854.79	14868.25	14873.40

Table 5: *GE for $X_1 \sim TCh$ and $X_2 \sim TCh$ and different upper truncation limits (standard errors between parentheses).*

	max	max×10	max×100
α_1	1.3344 (0.0950)	1.3420 (0.0489)	1.3423 (0.0492)
α_2	1.1767 (0.0444)	1.1771 (0.0431)	1.1706 (0.0427)
H_1	631.1100 (36.2700)	623.2490 (35.6333)	619.3690 (35.1012)
H_2	76.8340 (4.9617)	77.7100 (5.0423)	78.2220 (5.1128)
ω	3.0290×10^{-8} (1.4497×10^{-8})	2.3070×10^{-9} (2.6310×10^{-9})	1.9540×10^{-10} (8.773163×10^{-10})
$\ln L$	−7419.63	−7434.21	−7436.87
AIC	14849.26	14878.42	14883.74

Table 6: *PE for $X_1 \sim TCh$ and $X_2 \sim TCh$ and different upper truncation limits (standard errors between parentheses).*

	max	max×10	max×100
α_1	1.3362 (0.0497)	1.3409 (0.0492)	1.3407 (0.0492)
α_2	1.1564 (0.0755)	1.1693 (0.0768)	1.1706 (0.0769)
H_1	624.1119 (35.6350)	623.3819 (35.4695)	623.5835 (35.4896)
H_2	78.9157 (6.2698)	78.3094 (6.2351)	78.2899 (6.2332)
ω	9.4918×10^{-9} (9.6420×10^{-9})	9.7283×10^{-10} (3.2613×10^{-9})	9.7508×10^{-11} (2.4705×10^{-9})
$\ln L$	−7435.49	−7437.15	−7437.26
AIC	14880.99	14884.31	14884.52

Table 7: *VaR values for several truncated Sarmanov distributions and different confidence levels using GE.*

Distribution	Confidence level q			
$\phi_i = x - E[X_i]$	0.95	0.99	0.995	0.999
$Log - Sarmanov$ $TN + TNmixt(\max)$	3484.592	11221.492	16469.954	34770.477
$Log - Sarmanov$ $TN + TNmixt(\max \times 10)$	6703.136	18043.612	26888.181	62829.676
$Log - Sarmanov$ $TN + TNmixt(\max \times 100)$	6363.582	15658.461	21929.474	44858.422
$Sarmanov$ $TCh + TCh(\max)$	3307.784	16192.401	27588.607	71445.821
$Sarmanov$ $TCh + TCh(\max \times 10)$	6399.348	20755.411	34073.251	113319.114
$Sarmanov$ $TCh + TCh(\max \times 100)$	6405.983	20868.242	34416.052	106865.442
Empirical values	7905.600	24821.140	28420.870	92112.930

Table 8: *TVaR values for several truncated Sarmanov distributions and different confidence levels using GE.*

Distribution	Confidence level q			
$\phi_i = x - E[X_i]$	0.95	0.99	0.995	0.999
$Log - Sarmanov$ $TN + TNmixt(\max)$	10722.435	23966.182	31500.887	55354.994
$Log - Sarmanov$ $TN + TNmixt(\max \times 10)$	15198.865	35093.589	48500.570	88741.558
$Log - Sarmanov$ $TN + TNmixt(\max \times 100)$	13184.039	28457.449	38574.933	75462.460
$Sarmanov$ $TCh + TCh(\max)$	14236.400	40314.549	56765.029	103080.945
$Sarmanov$ $TCh + TCh(\max \times 10)$	20317.585	59231.604	92295.169	244553.953
$Sarmanov$ $TCh + TCh(\max \times 100)$	21255.717	63750.746	101169.957	284614.349
Empirical values	20836.960	49453.170	73078.330	149791.000

Table 9: *VaR values for several truncated Sarmanov distributions and different confidence levels using PE.*

Distribution	Confidence level q			
$\phi_i = x - E[X_i]$	0.95	0.99	0.995	0.999
$Log - Sarmanov$ $TN + TNmixt(\max)$	6146.651	15182.363	20345.663	36692.056
$Log - Sarmanov$ $TN + TNmixt(\max \times 10)$	6499.495	16546.371	22518.986	46979.657
$Log - Sarmanov$ $TN + TNmixt(\max \times 100)$	6485.068	16408.995	22269.078	36658.346
$Sarmanov$ $TCh + TCh(\max)$	5943.208	19699.385	29599.685	77228.707
$Sarmanov$ $TCh + TCh(\max \times 10)$	6229.109	23097.386	38009.139	116412.203
$Sarmanov$ $TCh + TCh(\max \times 100)$	6237.787	23074.898	38462.701	141907.139
Empirical values	7905.600	24821.140	28420.870	92112.930

Table 10: *TVaR values for several truncated Sarmanov distributions and different confidence levels using PE.*

Distribution	Confidence level q			
$\phi_i = x - E[X_i]$	0.95	0.99	0.995	0.999
$Log-Sarmanov$ $TN+TNmixt(\max)$	12100.52	24887.46	32197.80	57251.34
$Log-Sarmanov$ $TN+TNmixt(\max\times 10)$	14002.47	31249.10	42864.60	91912.35
$Log-Sarmanov$ $TN+TNmixt(\max\times 100)$	15727.76	39941.33	60676.94	190957.41
$Sarmanov$ $TCh+TCh(\max)$	16015.06	40962.05	57834.45	95298.44
$Sarmanov$ $TCh+TCh(\max\times 10)$	20355.52	58995.50	89891.05	191242.66
$Sarmanov$ $TCh+TCh(\max\times 100)$	21601.87	64854.65	101058.90	222460.19
Empirical values	20836.960	49453.170	73078.330	149791.000

Acknowledgements

Zuhair Bahraoui and Catalina Bolancé acknowledge the support received from Spanish Ministry and FEDER grant ECO2013-48326-C2-1-P. We also thank the members and affiliates of Riskcenter at the University of Barcelona. The authors gratefully acknowledge the two anonymous referees for very helpful comments and recommendations.

References

Artzner, P., Delbaen, F., Eber, J. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9, 203–228.

Bahraoui, Z., Bolancé, C. and Pérez-Marín, M. (2014). Testing extreme value copulas to estimate the quantile. *SORT-Statistics and Operations Research Transactions*, 38, 89–102.

Bairamov, I., Altinsoy, B. and Kerns, G. (2011). On generalized Sarmanov bivariate distributions. *TWMS Journal of Applied and Engineering Mathematics*, 1, 86–97.

Bairamov, I., Kotz, S. and Gebizlioglu, O. L. (2001). The Sarmanov family and its generalization: theory and methods. *South African Statistical Journal*, 35, 205–224.

Bolancé, C., Bahraoui, Z. and Artís, M. (2014). Quantifying the risk using copulae with nonparametric marginals. *Insurance: Mathematics and Economics*, 58, 64–56.

Bolancé, C., Guillén, M., Pelican, E. and Vernic, R. (2008). Skewed bivariate models and non-parametric estimation for CTE risk measure. *Insurance: Mathematics and Economics*, 43, 386–393.

Bourden, R. and Faires, J. (2001). *Numerical Analysis*. 7th Edition. Brooks/Cole, Pacific Grove, CA.

Buch-Larsen, T., Guillén, M., Nielsen, J. and Bolancé, C. (2005). Kernel density estimation for heavy-tailed distributions using the Champernowne transformation. *Statistics*, 39, 503–518.

Champernowne, D. (1952). The graduation of income distribution. *Econometrica*, 20, 591–615.

Cohen, L. (1984). Probability distributions with given multivariate marginals. *Journal of Mathematical Physics*, 25, 2402–2403.

Danaher, P. (2007). Modeling page views across multiple websites with an application to internet reach and frequency prediction. *Marketing Science*, 26, 422–437.

Farlie, D. (1960). The performance of some correlation coeficients for a general bivariate distribution. *Biometrika*, 47, 307–323.

Fisk, P. (1961). The graduation of income distribution. *Econometrica*, 29, 171–185.

Gómez-Déniz, E. and Pérez-Rodríguez, J. (2015). Closed form solution for a bivariate distribution in stochastic frontier models with dependent errors. *Journal of Productivity Analysis*, 43, 215–223.

Hashorva, E. and Ratovomirija, G. (2015). On Sarmanov mixed Erlang risks in insurance applications. *ASTIN Bulletin*, 45, 175–205.

Hernández-Bastida, A. and Fernández-Sánchez, M. (2013). A Sarmanov family with beta and gamma marginal distributions: an application to the Bayes premium in a collective risk model. *Statistical Methods & Applications*, 21, 391–409.

Hernández-Bastida, A., Fernández-Sánchez, M. and Gómez-Déniz, E. (2009). The net Bayes premium with dependence between the risk profiles. *Insurance: Mathematics and Economics*, 45, 247–254.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, New York.

Kotz, S., Balakrishnan, N. and Johnson, N. (2000). *Continuous Multivariate Distributions*. Vol.1: Models and Applications. John Wiley & Sons, Inc., New York.

Lee, M. (1996). Properties and applications of the Sarmanov family of bivariate distributions. *Communications in Statistics-Theory and Methods*, 25, 1207–1222.

Mladenovic, N. and Hansen, P. (1997). Variable neighborhood search. *Computers and Operations Research*, 24, 1097–1100.

Nelsen, R. (2006). *An Introduction to Copulas*. 2nd ed. Springer, New York.

Pelican, E. and Vernic, R. (2013). Maximum-likelihood estimation for the multivariate Sarmanov distribution: simulation study. *International Journal of Computer Mathematics*, 90, 1958–1970.

Sarmanov, O. (1966). Generalized normal correlation and two-dimensional Frechet classes. *Doclady (Soviet Mathematics)*, 168, 596–599.

Shubina, M. and Lee, T. (2004). On maximum attainable correlation and other measures of de- pendence for the Sarmanov family of bivariate distributions. *Communications in Statistics: Theory and Methods*, 33, 1031–1052.

Vernic, R. (2014). On the distribution of a sum of Sarmanov distributed random variables. *Journal of Theoretical Probability*, Forthcoming.

Yang, Y. and Hashorva, E. (2013). Extremes and products of multivariate AC-product risks. *Insurance: Mathematics and Economics*, 52, 312–319.