Insurance: Mathematics and Economics 71 (2016) 154–163

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Risk aggregation in multivariate dependent Pareto distributions

José María Sarabia [a,*], Emilio Gómez-Déniz [b], Faustino Prieto [a], Vanesa Jordá [a]

[a] Department of Economics, University of Cantabria, Avda de los Castros s/n, 39005-Santander, Spain

[b] Department of Quantitative Methods in Economics and TiDES Institute, University of Las Palmas de Gran Canaria, 35017-Las Palmas de G.C., Spain

ARTICLE INFO

Article history:
Received May 2015
Received in revised form
July 2016
Accepted 30 July 2016
Available online 24 August 2016

Keywords:
Dependent risks
Individual risk model
Collective risk model
Classical Pareto distribution
Hypergeometric functions

ABSTRACT

In this paper we obtain closed expressions for the probability distribution function of aggregated risks with multivariate dependent Pareto distributions. We work with the dependent multivariate Pareto type II proposed by Arnold (1983, 2015), which is widely used in insurance and risk analysis. We begin with an individual risk model, where the probability density function corresponds to a second kind beta distribution, obtaining the VaR, TVaR and several other tail risk measures. Then, we consider a collective risk model based on dependence, where several general properties are studied. We study in detail some relevant collective models with Poisson, negative binomial and logarithmic distributions as primary distributions. In the collective Pareto–Poisson model, the probability density function is a function of the Kummer confluent hypergeometric function, and the density of the Pareto–negative binomial is a function of the Gauss hypergeometric function. Using data based on one-year vehicle insurance policies taken out in 2004–2005 (Jong and Heller, 2008) we conclude that our collective dependent models outperform other collective models considered in the actuarial literature in terms of AIC and CAIC statistics.

1. Introduction

The individual and collective risk models (Kaas et al., 2001; Klugman et al., 2008 respectively) assume independence between: (i) different claim amounts; (ii) the number of claims and claim amounts and (iii) claim amounts and inter-claim times. This facilitates the computation of many risks measures, but can be restrictive in different contexts. Some recent research seeks to generalize both individual and collective classical models by considering some kind of dependence structure.

Sarabia and Guillén (2008) consider extensions of the classical collective model assuming that the conditional distributions $S|N$ and $N|S$ belong to some prescribed parametric family, where S is the total claim amount and N is the number of claims. Using conditional specification techniques (Gómez-Déniz and Calderín, 2014) have obtained discrete distributions to be used in the collective risk model to compute the right-tail probability of the aggregate claims size distribution.

Albrecher and Teugels (2006) consider a copula dependence structure for the interclaim time and the subsequent claim size. Boudreault et al. (2006) study an extension of the classical compound Poisson risk model, where the distribution of the next claim amount is a function of the time elapsed since the last claim. Cossette et al. (2008) consider another extension introducing a dependence structure between the claim amounts and the inter-claim time using a generalized Farlie–Gumbel–Morgenstern copula. Cossette et al. (2004) employ a variation of the compound binomial model in a Markovian environment, which is an extension of the model presented by Gerber (1988). Compound Poisson approximations for individual dependent risks are considered in Genest et al. (2003).

Finally, Cossette et al. (2013) consider a portfolio of dependent risks whose multivariate distribution is the Farlie–Gumbel–Morgenstern copula with mixed Erlang distribution marginals.

In this paper we obtain closed expressions for the probability distribution function of aggregated risks with multivariate dependent Pareto distributions between the different claim amounts. We work with the dependent multivariate Pareto type II proposed by Arnold (1983, 2015), which is widely used in insurance and risk analysis. In the classic individual risk model, we show that the probability density function (pdf) corresponds to a beta distribution of the second kind. Then we obtain several risk measures including the VaR and other tail risks measures. Next, we study the general properties of a collective model with dependent risks, focusing on some relevant collective models with Poisson, negative binomial and logarithmic distributions as primary distributions.

* Corresponding author.
E-mail addresses: sarabiaj@unican.es (J.M. Sarabia), emilio.gomez-deniz@ulpgc.es (E. Gómez-Déniz), faustino.priero@unican.es (F. Prieto), vanesa.jorda@unican.es (V. Jordá).

http://dx.doi.org/10.1016/j.insmatheco.2016.07.009

For these three models we obtain simple and closed expressions for the aggregated distributions.

The contents of this paper are the following: In Section 2 we present the main univariate distributions used in the paper; Section 3 examines the class of multivariate dependent Pareto distributions for modeling aggregated risks. Section 4 presents the individual risk model under dependence and Section 5 introduces the collective risk model under dependence. After presenting general results we study the compound models where the primary distribution is Poisson, negative binomial, geometric and logarithmic and the secondary distribution is Pareto. Section 6 includes an example with real data. The conclusions of the paper are given in Section 7.

2. Univariate distributions

In this section, we introduce several univariate random variables which will be used in the paper.

We work with the Pareto distribution with pdf given by,

$$f(x;\alpha,\beta)=\frac{\alpha}{\beta(1+x/\beta)^{\alpha+1}},\quad x>0, \tag{1}$$

and $f(x;\alpha,\beta)=0$ if $x<0$, where $\alpha,\beta>0$. Here, α is a shape parameter and β is a scale parameter. We represent $X\sim\mathcal{P}a(\alpha,\beta)$.

We denote by $X\sim\mathcal{G}a(\alpha)$ a gamma random variable with pdf $f(x)=\frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}$ if $x>0$, with $\alpha>0$. The exponential distribution with mathematical expectation 1 is denoted by $\mathcal{G}a(1)$.

The following lemma provides a simple stochastic representation of the Pareto distribution as quotient of random variables. The proof is straightforward and will be omitted.

Lemma 1. *Let U_1 and U_α independent gamma random variables such that $U_1\sim\mathcal{G}a(1)$ and $U_\alpha\sim\mathcal{G}a(\alpha)$, where $\alpha>0$. If $\beta>0$, the random variable,*

$$X=\beta\frac{U_1}{U_\alpha}\sim\mathcal{P}a(\alpha,\beta). \tag{2}$$

An extension of the Pareto distribution (1) is the following. A random variable X is said to be a beta distribution of the second kind if its pdf is of the form,

$$f(x;p,q,\beta)=\frac{x^{p-1}}{\beta^pB(p,q)(1+x/\beta)^{p+q}},\quad x>0, \tag{3}$$

and $f(x;p,q,\beta)=0$ if $x<0$, where $p,q,\beta>0$ and $B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ denotes the beta function. This random variable corresponds to the Pearson VI distribution in the classical Pearson systems of distributions and we write $X\sim\mathcal{B}2(p,q,\lambda)$. If we set $p=1$ in (3), we obtain a Pareto distribution $\mathcal{P}a(q,\beta)$ like (1).

The beta distribution of the second kind has a simple stochastic representation as a ratio of gamma random variables. As a direct extension of Lemma 1, if U_p and U_q are independent gamma random variables, the new random variable $X=\beta\frac{U_p}{U_q}$ has the pdf defined in (3).

3. The multivariate Pareto class

Now we present the class of multivariate dependent Pareto distribution which will be used in the different models.

In the literature several classes of multivariate Pareto distributions have been proposed. One of the main classes was introduced by Arnold (1983, 2015), in the context of the hierarchy Pareto distributions proposed by this author. Other classes were proposed by Chiragiev and Landsman (2009) and Asimit et al. (2010). The conditional dependence structure is the base of the construction of the proposals by Arnold (1987) and Arnold et al. (1993) (see also Arnold et al., 2001), where two different dependent classes are obtained.

Definition 1. Let $Y_1,Y_2,\ldots,Y_n$ and Y_α be mutually independent gamma random variables with distributions $Y_i\sim\mathcal{G}a(1)$, $i=1,2,\ldots,n$ and $Y_\alpha\sim\mathcal{G}a(\alpha)$ with $\alpha>0$. The multivariate dependent Pareto distribution is defined by the stochastic representation,

$$\boldsymbol{X}=(X_1,X_2,\ldots,X_n)^\top=\left(\beta\frac{Y_1}{Y_\alpha},\beta\frac{Y_2}{Y_\alpha},\ldots,\beta\frac{Y_n}{Y_\alpha}\right)^\top, \tag{4}$$

where $\beta>0$.

Note that the common random variable Y_α introduces the dependence in the model.

3.1. Properties of the multivariate Pareto class

We describe several properties of the multivariate Pareto defined in (4).

- Marginal distributions. By construction, the marginal distributions are Pareto,

  $X_i\sim\mathcal{P}a(\alpha,\beta),\quad i=1,2,\ldots,n.$
- The joint pdf of the vector $\boldsymbol{X}$ is given by,

$$f(x_1,\ldots,x_n;\alpha,n)=\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\beta^n}\frac{1}{\left(1+\sum_{i=1}^n x_i/\beta\right)^{\alpha+n}},$$

$$x_1,\ldots,x_n>0. \tag{5}$$

  This expression corresponds to the joint pdf of the multivariate Pareto type II proposed by Arnold (1983, 2015).
- The covariance is given by,

$$cov(X_i,X_j)=\frac{\beta^2}{(\alpha-1)^2(\alpha-2)},\quad \alpha>2,\ i\neq j$$

  and the correlation between components is,

$$\rho(X_i,X_j)=\frac{1}{\alpha},\quad \alpha>2,\ i\neq j.$$
- General moments. The moments of (1) are,

$$E(X_1^{r_1}\cdots X_n^{r_n})=\frac{\Gamma(\alpha-A)}{\Gamma(\alpha)}\prod_{i=1}^n\beta^{r_i}\Gamma(1+r_i),$$

  where $A=r_1+\cdots+r_m$ and $\alpha>A$.

  The dependence structure of $\boldsymbol{X}$ is studied in the following result

Proposition 1. *The random variables $\boldsymbol{X}=(X_1,\ldots,X_n)^\top$ are associated, and then $cov(X_i,X_j)\geq 0$, if $i\neq j$.*

Proof. See Appendix. ■

Remark. Let us consider the multivariate Pareto survival function of (5) given by,

$$\bar{F}(x_1,\ldots,x_n)=\left(1+\sum_{i=1}^n\frac{x_i}{\beta}\right)^{-\alpha},\quad x_1,\ldots,x_n>0,$$

with $\alpha,\beta>0$. For this family, the associated copula is the Pareto copula or Clayton copula,

$$C(u_1,\ldots,u_n;\alpha)=\left(u_1^{-1/\alpha}+\cdots+u_n^{-1/\alpha}-n+1\right)^{-\alpha}.$$

Note that the dependence increases with α, being the independence case obtained when $\alpha\to 0$ and the Fréchet upper bound when $\alpha\to\infty$.

4. The individual risk model under Pareto dependence

In this section we consider the individual risk model assuming dependence between risks. The distribution of sums of i.i.d. Pareto distributions was obtained by Ramsay (2006). Let $(X_1, \ldots, X_n)^\top$ be the multivariate Pareto distribution defined in (4). Then, we consider the aggregate risks $S_n = X_1 + \cdots + X_n$. We have the following result.

Theorem 1. *The pdf of the aggregate random variable S_n, where the components are Pareto defined in (1) is given by,*

$$f_{S_n}(x; n, \alpha, \beta) = \frac{x^{n-1}}{\beta^n B(n, \alpha)(1 + x/\beta)^{n+\alpha}}, \quad x > 0 \tag{6}$$

and $f_{S_n}(x; n, \alpha, \beta) = 0$ if $x < 0$, that is $S_n \sim \mathcal{B}2(n, \alpha, \beta)$.

Proof. See Appendix. ■

Note that (6) corresponds to the pdf of a second kind beta distribution defined in (3). Fig. 1 represents the pdf (6) for $\alpha = 1/2$; 1; 2 and 10 for $n = 2, 5, 10$ and 20.

4.1. Risk measures

Here we present some risk measures for the second kind beta distribution, which can be applied for the aggregate pdf given in (6). The value at risk VaR at level u, with $0 < u < 1$ of a random variable X with cumulative distribution function (cdf) $F(x)$ is defined as,

$$\mathrm{VaR}[X; u] = \inf\{x \in \mathbb{R},\ F(x) \geq u\}.$$

If $X \sim \mathcal{B}2(p, q, \beta)$, Guillén et al. (2013) have obtained that,

$$\mathrm{VaR}[X; u] = \beta \frac{IB^{-1}(u; p, q)}{1 - IB^{-1}(u; p, q)}, \tag{7}$$

where $IB^{-1}(u; p, q)$ denotes the inverse of the incomplete ratio beta function, which corresponds to the quantile function of the classical beta distribution of the first kind. Then, using (7) for the aggregate distribution (6) we have

$$\mathrm{VaR}[S_n; u] = \beta \frac{IB^{-1}(u; n, \alpha)}{1 - IB^{-1}(u; n, \alpha)},$$

with $0 < u < 1$.

The following result provides higher moments of the worst x_u events, which extend popular risk measures.

Lemma 2. *Let $X \sim \mathcal{B}2(p, q, \beta)$ be a second kind beta distribution. Then, the conditional tail moments are given by ($q > r$),*

$$E(X^r | X > x_u) = \frac{\beta^r \Gamma(p + r)\Gamma(q - r)}{(1 - u)\Gamma(p)\Gamma(q)} \times \left\{1 - B\left(\frac{x_u/\beta}{1 + x_u/\beta}; p + r, q - r\right)\right\}, \tag{8}$$

where $x_u = VaR[X; u]$ represents the value at risk with $u \in (0, 1)$.

Proof. See Appendix. ■

If we take $r = 1$ in (8) we obtain the tail value at risk TVaR, which was obtained by Guillén et al. (2013).

5. The collective risk model under dependence

In this section we consider the collective model based on dependence between claim amounts. Let N be the number of claims in a portfolio of policies in a time period. Let X_i, $i = 1, 2, \ldots$ be the amount of the ith claim and $S_N = X_1 + \cdots + X_N$ the aggregate claims of the portfolio in the time period considered.

5.1. General properties

We consider two assumptions: (1) we assume that all the claims $X_1, X_2, \ldots, X_n, \ldots$ are dependent random variables with the same distribution and (2) the random variable N is independent of all claims $X_1, X_2, \ldots, X_n, \ldots$

Theorem 2. *Let $X_1, X_2, \ldots, X_n, \ldots$ be dependent random variables with common cdf $F(x)$, and let N be the observed number of claims, with probability mass function (pmf) $p_n = \Pr(N = n)$, for $n = 0, 1, \ldots$, which is independent of all X_i's, $i = 1, 2, \ldots$ Then, the cdf of the aggregate losses S_N is,*

$$F_{S_N}(x) = \sum_{n=0}^{\infty} p_n F_X^{(n)}(x),$$

where $F_X^{(n)}(x)$ represent the cdf of the convolution of the n dependent claims $(X_1, \ldots, X_n)$.

Proof. See Appendix. ■

The mean and the variance of the collective model can be found in the following result.

Lemma 3. *The mean and the variance of S_N under dependence are given by,*

$$E(S_N) = E(N)E(X), \tag{9}$$

$$var(S_N) = E(N)var(X) + var(N)(E(X))^2 + E[N(N-1)]cov(X_i, X_j). \tag{10}$$

Proof. See Appendix. ■

If the claims X_i and X_j are independent $cov(X_i, X_j) = 0$ and then (10) becomes the usual formula for $var(S_N)$.

On the other hand, if the random variables $\{X_i\}$ are associated (see Appendix),

$$\mathrm{var}(S_N^{(I)}) < \mathrm{var}(S_N),$$

where $\mathrm{var}(S_n^{(I)})$ is the variance in the independent case and $\mathrm{var}(S_n)$ the variance in the dependent case given by Eq. (10). This fact is a consequence of the associated property, which leads to a positive covariance between X_i and X_j.

5.2. Compound Pareto models

In this section we obtain the distribution of the compound collective model $S_N = X_1 + X_2 + \cdots + X_N$ where the secondary distribution is Pareto given by (4), and several primary distribution are considered for N.

In the case of independence between claims X_i this distribution was studied by Ramsay (2009) when the primary distribution is Poisson and negative binomial cases.

5.2.1. The compound Pareto–Poisson distribution

We consider the model where the primary distribution is a Poisson distribution.

Theorem 3. *If we assume a Poisson distribution with parameter λ as primary distribution and $(X_1, \ldots, X_n)$ is defined in (4), the pdf of the random variable S_N is given by,*

$$f_{S_N}(x; \alpha, \lambda, \beta) = \frac{\alpha\lambda e^{-\lambda}}{\beta(1 + x/\beta)^{\alpha+1}} {}_1F_1\left[1 + \alpha; 2; \frac{\lambda x/\beta}{1 + x/\beta}\right], \quad x > 0 \tag{11}$$

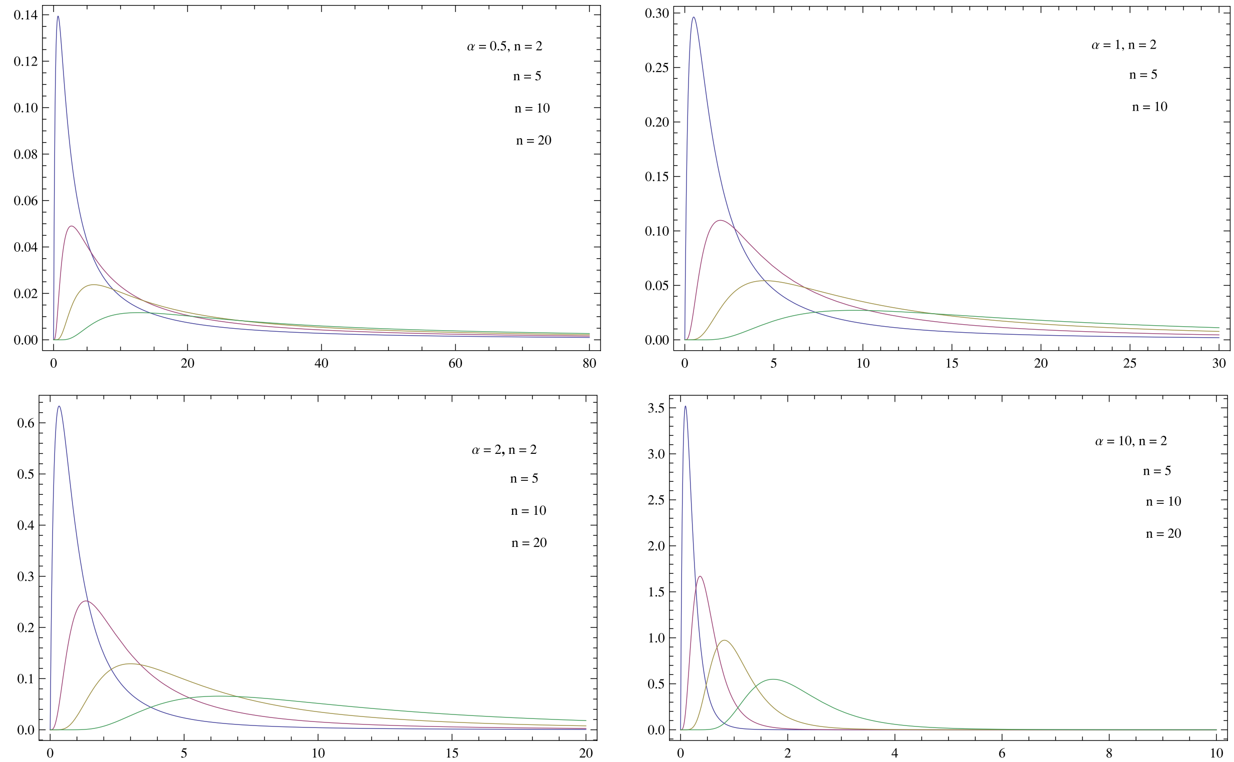

Fig. 1. Probability density function of S_n (Eq. (6)) the individual model for some selected values of α and n.

and $f_{S_N}(0; \alpha, \lambda, \beta) = e^{-\lambda}$ *where* ${}_1F_1(a; b; z)$ *denotes the Kummer confluent hypergeometric function defined by,*

$$ {}_1F_1[a; b; z] = \sum_{n=0}^{\infty} \frac{(a)_n z^n}{(b)_n n!}, $$

where $(a)_n$ *represent the Pochhammer symbol defined by* $(a)_n = a(a-1)\dots(a-n+1)$.

Proof. See Appendix. ■

Fig. 2 represents the pdf (11) Pareto–Poisson for some values of $\alpha = 0, 5; 1; 2$ and 10 and $\lambda = 2, 5, 10$ and 20, taking $\beta = 1$.

Using formulas (9) and (10), the mean and variance of the dependent Pareto–Poisson collective model are given by,

$$ E(S_N) = \frac{\lambda\beta}{\alpha-1}, \quad \alpha > 1, $$
$$ var(S_N) = \frac{\alpha\beta^2(\lambda+2\alpha-2)}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2. $$

5.2.2. The compound Pareto–negative binomial distribution

Let N be a negative binomial distribution with pmf,

$$ \Pr(N=n) = \frac{\Gamma(n+r)}{\Gamma(n+1)\Gamma(r)} p^r (1-p)^n, \quad n = 0, 1, 2, \dots. \tag{12} $$

We have the following theorem.

Theorem 4. *If we assume a negative binomial distribution with parameter r and p and pmf given by (12) as primary distribution and* $(X_1, \dots, X_n)$ *is defined in (4), the pdf of the random variable* S_N *is given by,*

$$ f_{S_N}(x; \alpha, \beta, r, p) = \frac{r\alpha(1-p)p^r}{\beta(1+x/\beta)^{\alpha+1}} \times {}_2F_1\left[1+r, 1+\alpha; 2; \frac{(1-p)x/\beta}{1+x/\beta}\right], \quad x > 0 $$

and $f_{S_N}(0; \alpha, \beta, r, p) = p^r$, *where* ${}_2F_1(a, b; c; z)$ *denotes the Gauss hypergeometric function defined by,*

$$ {}_2F_1[a, b; c; z] = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n z^n}{(c)_n n!}, $$

where $(a)_n$ *denotes the Pochhammer symbol.*

The proof is similar to the proof in Theorem 3 and will be omitted.

Again, using formulas (9) and (10), the mean and variance of the dependent Pareto–negative binomial collective model are given by,

$$ E(S_N) = \frac{r(1-p)\beta}{p(\alpha-1)}, \quad \alpha > 1, $$
$$ var(S_N) = \frac{r(1-p)\beta^2((1+p)(\alpha-1)+r(1-p))}{p^2(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2. $$

5.2.3. The compound Pareto–geometric distribution

In this result we obtain the pdf of the Pareto–geometric distribution.

Corollary 1. *If we assume a geometric distribution with parameter and pmf given by* $\Pr(N=n) = p(1-p)^n$, $n = 0, 1, \dots$ *as primary distribution and* $(X_1, \dots, X_n)$ *is defined in (4), the pdf of the random variable* S_N *is given by,*

$$ f_{S_N}(x; \alpha, \beta, p) = \frac{\alpha p(1-p)}{\beta(1+px/\beta)^{\alpha+1}}, \quad x > 0 \tag{13} $$

and $f_{S_N}(0; \alpha, \beta, p) = p$.

Fig. 3 represents the pdf (13) for some selected values of α and p $(\beta = 1)$.

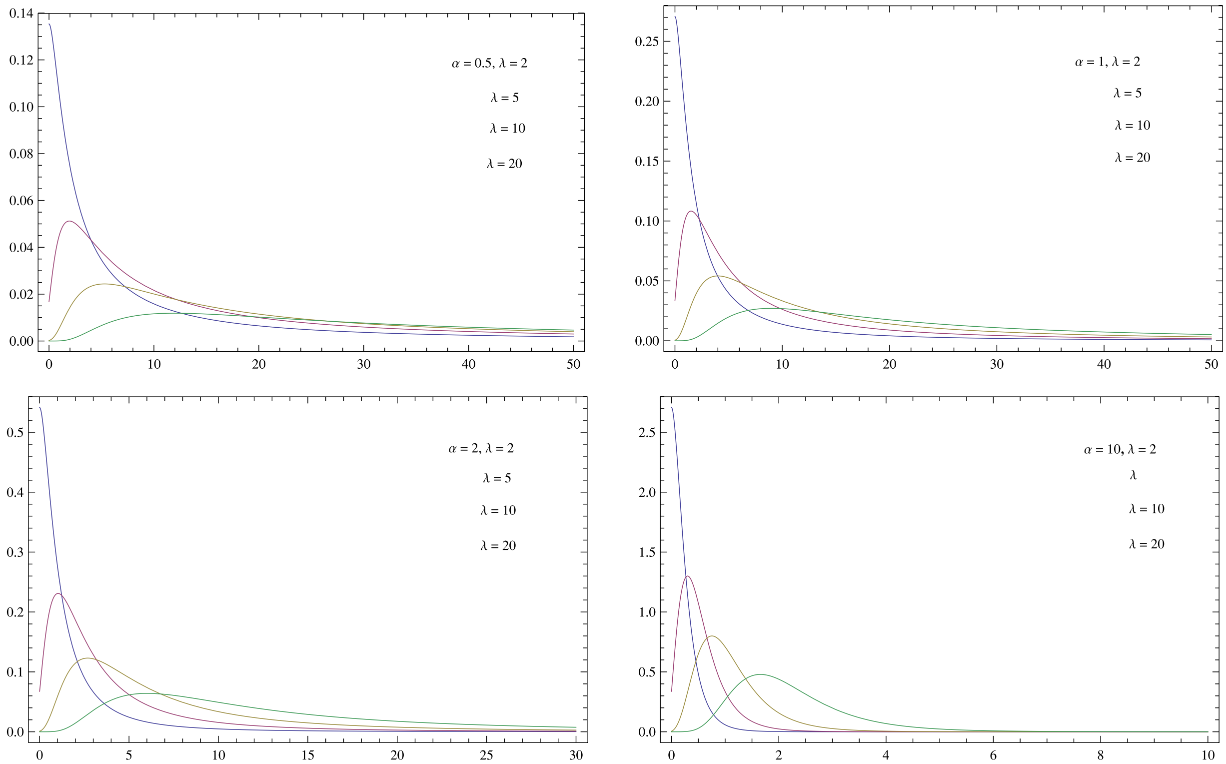

Fig. 2. Probability density function of the collective model Pareto–Poisson (Eq. (11)) for $\beta = 1$ and selected values of α and λ.

5.2.4. The compound Pareto–logarithmic distribution

Let N be a discrete logarithmic distribution with pmf,

$$\Pr(N = n) = -\frac{1}{\log(1-\theta)}\frac{\theta^n}{n}, \quad n = 1, 2, \ldots, \tag{14}$$

where $\theta \in (0, 1)$.

We have the following theorem.

Theorem 5. *If we assume a logarithmic distribution with parameter θ and pmf given by (14) as primary distribution and $(X_1, \ldots, X_n)$ is defined in (4), the pdf of the random variable S_N is given by,*

$$f_{S_N}(x; \alpha, \beta, \theta) = -\frac{1}{\log(1-\theta)}\left[\frac{1}{x(1+(1-\theta)x/\beta)^{\alpha}} - \frac{1}{x(1+x/\beta)^{\alpha}}\right], \quad x > 0 \tag{15}$$

and $f_{S_N}(x; \alpha, \beta, \theta) = 0$, if $x < 0$.

The proof of this result is omitted.

Fig. 4 represents the pdf (15) Pareto–Logarithmic for some values of $\alpha = 0, 5$; 1; 2 and 10 and $\theta = 0.2$; 0.4; 0.6 and 0.8 taking $\beta = 1$.

The mean and variance of the dependent Pareto–Logarithmic collective model are

$$E(S_N) = \frac{a\beta\theta}{(\alpha-1)(1-\theta)}, \quad \alpha > 1,$$

$$var(S_N) = \frac{a\beta^2\theta 2(\alpha-1) - \alpha\theta(1+a) + \theta(1+2a)}{(\alpha-1)^2(\alpha-2)(1-\theta)^2}, \quad \alpha > 2,$$

being $a = -1/\log(1-\theta)$.

6. A numerical application with real data

To compare the performance of the models presented in this paper, we examine data on one-year vehicle insurance policies taken out in 2004 or 2005. This data set is available on the website of the Faculty of Business and Economics, Macquarie University (Sydney, Australia) (see also Jong and Heller, 2008). The first 100 observations of this data set are shown in Table 1, with the following elements: from left to right, the policy number, the number of claims and the size of the claims. The total portfolio contains 67 856 policies of which 4624 have at least one claim. Some descriptive statistics are shown in Table 2. The standard deviation is very large for the size of the claims, which means that a premium based only on the mean claim size is not adequate for computing the bonus–malus premiums. The covariance between the claims and sizes is positive and takes the value 141.574.

Fig. 6 shows the complete number of claims and the total claim amount concerning these claims. It can be seen that the largest claim values appear in the case of single claims, while these values fall with larger numbers of claims (see Fig. 5).

To capture this, the compound Poisson model has been traditionally considered when the size of a single claim is modeled by an exponential distribution, chiefly because of the complexity of the collective risk model under other probability distributions such as Pareto and log-normal distributions.

Perhaps the most well-known aggregate claims model is obtained when the primary and secondary distributions are the Poisson and the exponential distributions, respectively. In this case, Rolski et al. (1999) and others, show that the distribution of the random variable total claim amount is given by

$$f_S(x; \alpha, \lambda) = \sqrt{\frac{\lambda\alpha}{x}}\exp(-\lambda-\alpha x) I_1\left(2\sqrt{\lambda\alpha x}\right), \quad x > 0,$$

while $f_S(0; \alpha, \lambda) = \exp(-\lambda)$. Here, $\lambda > 0$ and $\alpha > 0$ are the parameters of the Poisson and exponential distributions, respectively and

$$I_\nu(z) = \sum_{k=0}^{\infty}\frac{(z/2)^{2k+\nu}}{\Gamma(k+1)\Gamma(\nu+k+1)}, \quad z \in \mathbb{R}, \ \nu \in \mathbb{R}, \tag{16}$$

represents the modified Bessel function of the first kind.

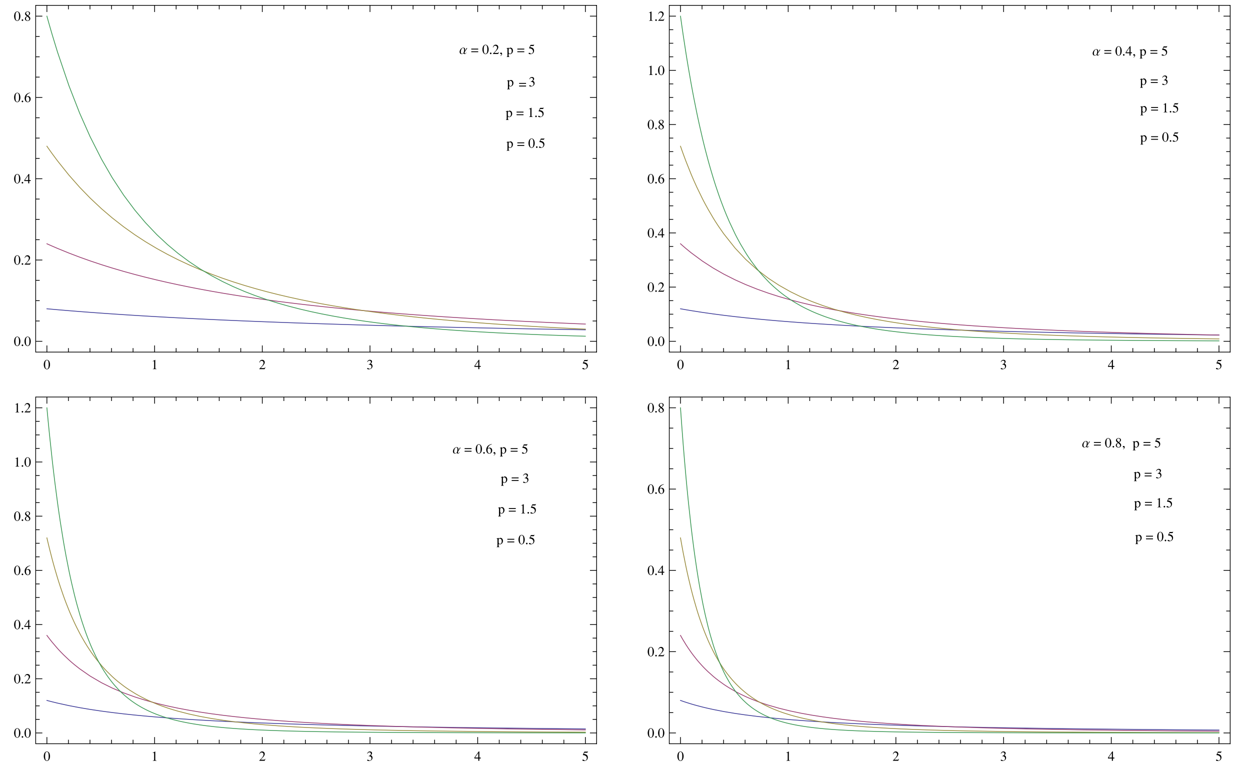

Fig. 3. Probability density function of the collective model Pareto–Geometric (Eq. (13)) for $\beta = 1$ and selected values of α and p.

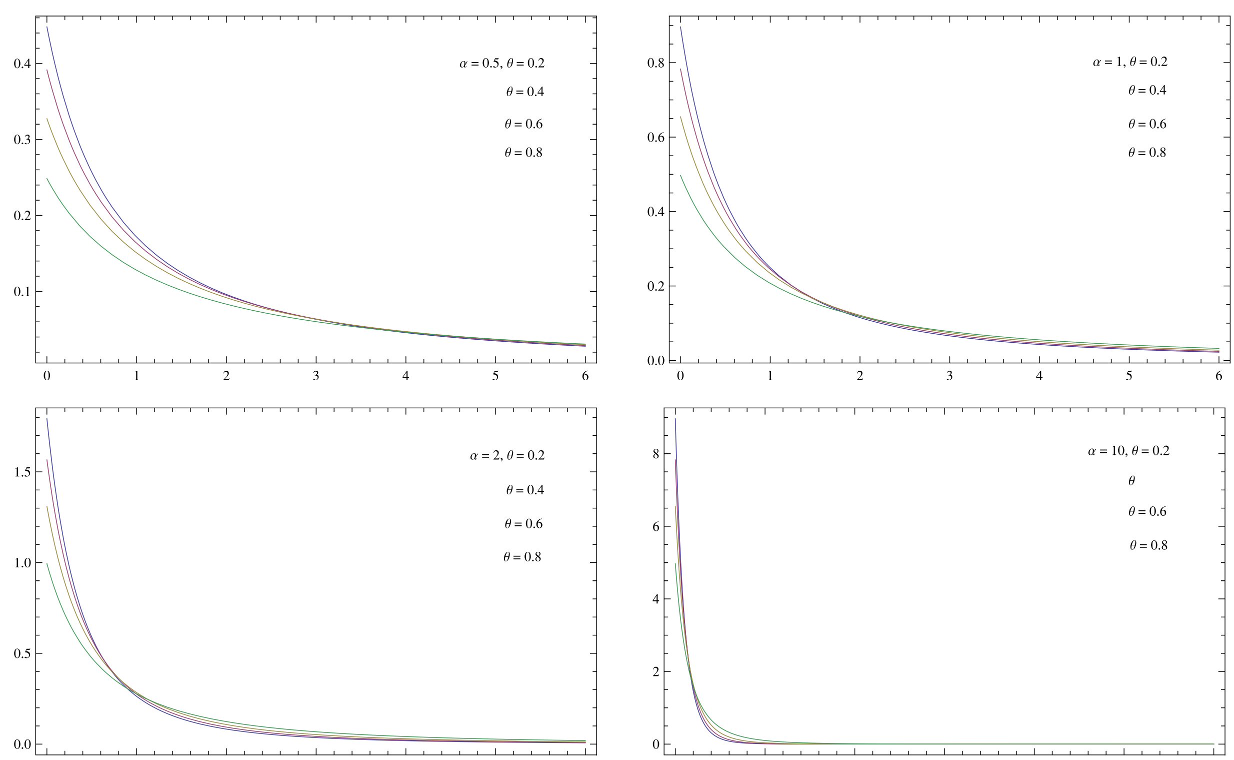

Fig. 4. Probability density function of the collective model Pareto–Logarithmic (Eq. (15)) for $\beta = 1$ and selected values of α and θ.

Additionally, the negative binomial distribution with parameters $r > 0$ and $0 < p < 1$ could also be assumed as primary distribution and the exponential distribution as the secondary distribution. In this case, the pdf of the random variable total claim amount (see Rolski et al., 1999) is now given by the expression

$$f_s(x; \alpha, r, p) = \alpha r p^r (1-p) \exp(-\alpha x) \times {}_1F_1[1+r; 2; \alpha(1-p)x], \quad x > 0,$$

Table 1
First 100 observations of the data set considered. From left to right: the policy number; the number of claims and the size of the claims.

1	0	0	21	0	0	41	2	1811.71	61	0	0	81	0	0
2	0	0	22	0	0	42	0	0	62	0	0	82	0	0
3	0	0	23	0	0	43	0	0	63	0	0	83	0	0
4	0	0	24	0	0	44	0	0	64	0	0	84	0	0
5	0	0	25	0	0	45	0	0	65	1	5434.44	85	0	0
6	0	0	26	0	0	46	0	0	66	1	865.79	86	0	0
7	0	0	27	0	0	47	0	0	67	0	0	87	0	0
8	0	0	28	0	0	48	0	0	68	0	0	88	0	0
9	0	0	29	0	0	49	0	0	69	0	0	89	0	0
10	0	0	30	0	0	50	0	0	70	0	0	90	0	0
11	0	0	31	0	0	51	0	0	71	0	0	91	0	0
12	0	0	32	0	0	52	0	0	72	0	0	92	0	0
13	0	0	33	0	0	53	0	0	73	0	0	93	0	0
14	0	0	34	0	0	54	0	0	74	0	0	94	0	0
15	1	669.51	35	0	0	55	0	0	75	0	0	95	0	0
16	0	0	36	0	0	56	0	0	76	0	0	96	1	1105.77
17	1	806.61	37	0	0	57	0	0	77	0	0	97	0	0
18	1	401.80	38	0	0	58	0	0	78	0	0	98	0	0
19	0	0	39	0	0	59	0	0	79	0	0	99	1	200
20	0	0	40	0	0	60	0	0	80	0	0	100	0	0

Table 2
Some descriptive data of claims and claim size for the data set.

	Number of claims	Total claim amount
Mean	0.072	137.27
Standard deviation	0.278	1056.30
min	0	0
max	4	55 922.10

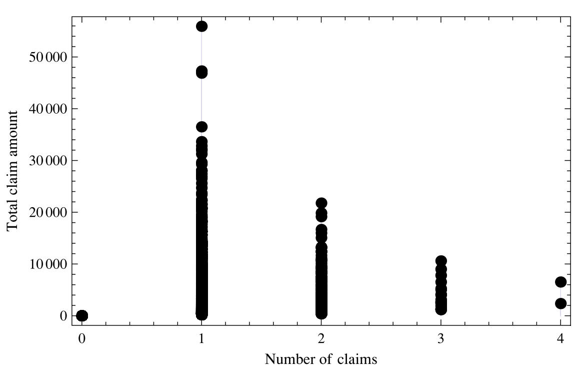

Fig. 5. Number of claims and their sizes.

where ${}_1F_1[\cdot;\cdot;\cdot]$ is the confluent hypergeometric function and $f_S(0;\alpha,r,p)=p^r$.

When $r=1$ in previous equation we get the pdf of the total claim amount when the geometric is the primary and the exponential is the secondary distribution. This results in

$$f_S(x;\alpha,p)=\alpha p(1-p)\exp(-\alpha p x),\quad x>0,$$

with $f_S(0;\alpha,p)=p$.

Finally, we consider the model for which the primary distribution is the mixture of the negative binomial distribution with parameters β and $\beta/(\beta+\mu)$, $\beta>0$, $\mu>0$ and where the pdf of the mean parameter μ, among policyholders in the insurance portfolio, is assumed to be a Pareto distribution with pdf given in (1). The secondary distribution is taken again as the exponential distribution with parameter $\lambda>0$. A more general mixed negative binomial–Pareto model has been considered in the actuarial literature by Shengwang et al. (1999) by assuming a generalized Pareto distribution with three parameters as mixing distribution. Recall that the sum of n independent and identically distributed exponential variables with parameter $\lambda>0$ is a gamma distribution with shape parameter n and scale parameter λ. Hence, after some computations, the pdf of the random variable $S_N=X_1+\cdots+X_N$, the total claim amount, is given by,

$$f_S(x;\alpha,\beta,\lambda)=\frac{\alpha\beta\lambda\exp(-\lambda x)\,{}_1F_1(\beta+1;\alpha+\beta+2;\lambda x)}{(\alpha+\beta+1)(\alpha+\beta)},\quad x>0,$$

while $f_S(0;\alpha,\beta)=\alpha/(\alpha+\beta)$.

Moment estimators can be obtained by equating the sample moments to the population moments. Furthermore, the parameters of the different models of the total claim amount can be estimated via maximum likelihood. We only present the case of Pareto–geometric case. To do so, consider a random sample $\{x_1,x_2,\dots,x_n\}$. The log-likelihood function can be written as

$$\begin{aligned}\ell(\alpha,\beta,p)&=n_0\log p+(n-n_0)\left[\log\alpha+\log p\right.\\&\quad\left.+\log(1-p)+\alpha\log\beta\right]\\&\quad-(\alpha+1)\sum_{x_i>0}\log(\beta+px_i),\end{aligned}$$

where n_0 is the number of zero-observations and $n-n_0$ is the number of non-zero sample observations, and n is the sample size.

In order to simplify the computations the values of the total claim amounts have been divided by 1000. The equations from which we get the maximum likelihood estimates cannot be solved explicitly. They must be solved either numerically or, as here, by directly maximizing the log-likelihood function. Since the global maximum of the log-likelihood surface is not guaranteed, different initial values in the parameter space were considered as a seed point. We use the `FindMaximum` function of Mathematica software package v.10.0 (Wolfram, 2003). Different maximization algorithms such as `Newton`, `PrincipalAxis` and `QuasiNewton` were used to ensure that our results are robust. Finally, the variances of the maximum likelihood estimates are the diagonal elements of the inverse matrix of negative second derivatives of the log-likelihood function, evaluated at the maximum likelihood estimates. When hypergeometric functions appear, these are replaced by their series representations by taking one hundred terms in the sum. This facilitates the computation of the Hessian matrix.

A summary of the results obtained is shown in Table 3. In this Table the estimated parameter values are presented together with their standard errors in parenthesis, the Akaike Information Criteria (AIC) and the Consistent Akaike Information Criteria (CAIC). Bozdogan (1987) proposed a corrected version of AIC in an attempt to overcome the tendency of the AIC to overestimate the complexity of the underlying model. Bozdogan (1987) observed that Akaike Information Criteria (AIC) (see Akaike, 1973) does not directly depend on sample size and as a result lacks certain properties of asymptotic consistency. In formulating CAIC, a correction

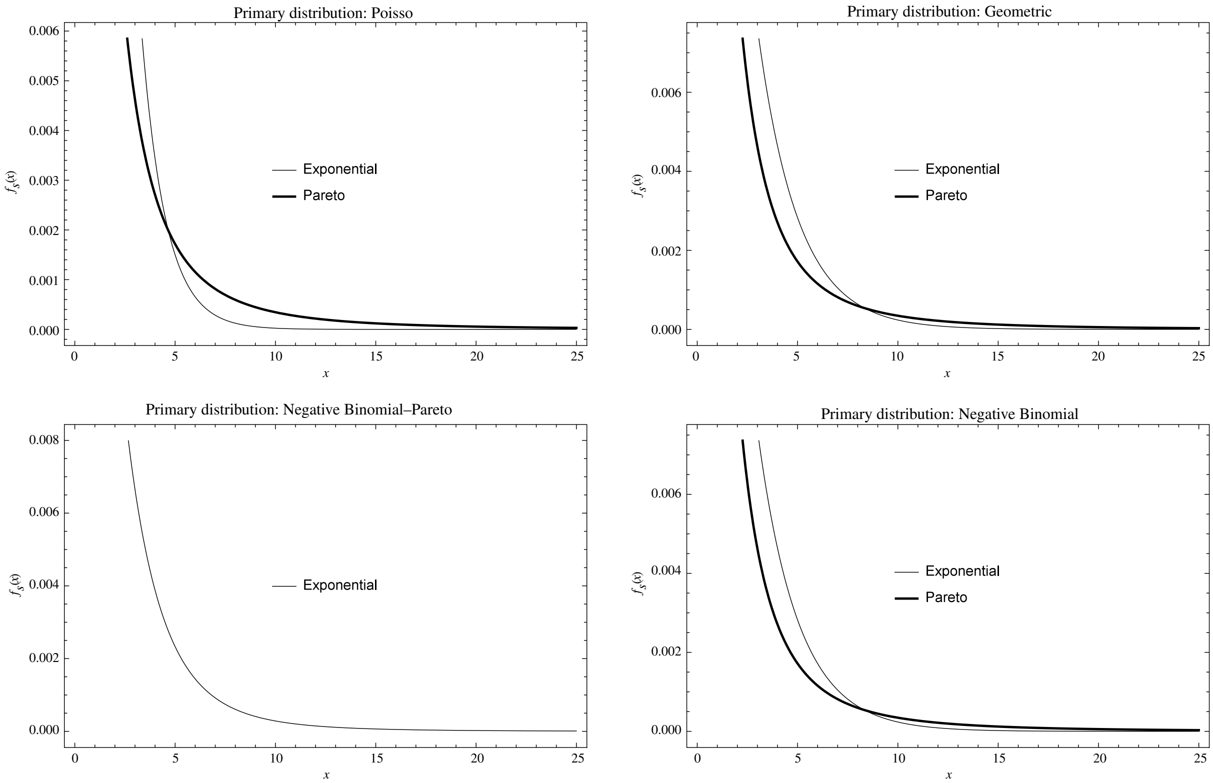

Fig. 6. Probability density function of the compound distributions for the estimated parameters of the models.

Table 3
Parameter estimates, standard errors (in parenthesis) and AIC and CAIC statistics for the different collective risk models considered, using data based on one-year vehicle insurance policies.

Distribution		$\widehat{r}$	$\widehat{p}$	$\widehat{\lambda}$	$\widehat{\alpha}$	$\widehat{\beta}$	AIC	CAIC
Primary	Secondary							
Poisson	Exponential			0.12057 (0.00104)	0.87832 (0.00759)		51 402.60	51 422.80
Geometric	Exponential		0.93186 (0.00097)	0.53273 (0.00785)			49 495.40	49 515.60
Negative binomial	Exponential	0.51168 (0.00000)	0.87090 (0.0000)		0.55250 (0.00471)		49 487.20	49 517.60
Neg.Bin.–Pareto	Exponential			0.82117 (0.01598)	2.84501 (0.12291)	0.20723 (0.00958)	49 010.00	49 040.40
Poisson	Pareto			0.07058 (0.00102)	2.04828 (0.00974)	2.13071 (0.04879)	48 229.50	48 259.90
Geometric	Pareto		0.93186 (0.00097)		2.04655 (0.08828)	2.05481 (0.12407)	48 229.60	48 260.00
Negative binomial	Pareto	0.31749 (0.21505)	0.80067 (0.12052)		2.05542 (0.09000)	1.91539 (0.17324)	48 232.10	48 272.60

factor based on the sample size is employed to compensate for the overestimating nature of AIC. The CAIC is defined as

$$\text{CAIC} = -2\log\ell + (1+\log n)\,k,$$

where again ℓ and k refer to the likelihood under the fitted model and the number of parameters, respectively and n is the sample size. As we can see, the AIC differs from CAIC in the second term which in this case takes into account the sample size n. Again, models that minimize the CAIC are selected. Our results indicate that the compound Pareto model outperforms the Poisson–exponential, the geometric–exponential, the negative binomial–exponential and the negative binomial–Pareto models. All these models, except the last one, have been widely considered in the actuarial literature when parametric specifications are used.

Fig. 6 shows the pdf of the four models considered using the parameter estimated given in Table 3. As we can see, the new Pareto compound models have a larger right tail than the traditional models based on the use of the exponential distribution.

Parameter estimates presented in Table 3 have been used to calculate the right-tail cumulative probabilities for different values of x as displayed in Table 4. As can be inferred from this Table for values of $x < 5$, the compound exponential–Poisson and exponential–geometric models has a slightly better performance than the new compound models in terms of the decreasing probabilities they generate. Nevertheless, the opposite result is obtained for $x > 5$. Similar results are obtained for the other models. Furthermore, new compound models tend to zero more slowly than the models based on the exponential secondary distribution.

7. Conclusions

We have obtained closed expressions for the distribution of aggregated risks with multivariate dependent Pareto type II distributions.

Table 4
Right-tail cumulative probabilities of the aggregate claim distribution for the estimated parameters of the models. The primary distribution above and the secondary below.

x	Poisson		Geometric		Neg. Bin.–Pareto	Neg. Bin.	
	Exponential	Pareto	Exponential	Pareto	Exponential	Exponential	Pareto
1	0.0496829	0.0317014	0.0414796	0.0316985	0.0387018	0.0415048	0.0317054
2	0.0217140	0.0181808	0.0252488	0.0181835	0.0227504	0.0252360	0.0181934
3	0.0094826	0.0117457	0.0153690	0.0117506	0.0138235	0.0153486	0.0117580
4	0.0041379	0.0081956	0.0093551	0.0082009	0.0086951	0.0093377	0.0082053
5	0.0018043	0.0060350	0.0056945	0.0060403	0.0056644	0.0056824	0.0060423
6	0.0007862	0.0046246	0.0034662	0.0046296	0.0038194	0.0034589	0.0046298
7	0.0003423	0.0036541	0.0021099	0.0036587	0.0026615	0.0021060	0.0036577
8	0.0001490	0.0029583	0.0012843	0.0029625	0.0019121	0.0012826	0.0029607
9	0.0000648	0.0024428	0.0007817	0.0024467	0.0014125	0.0007814	0.0024443
10	0.0000281	0.0020504	0.0004758	0.0020540	0.0010698	0.0004761	0.0020513
11	0.0000122	0.0017450	0.0002896	0.0017482	0.0008284	0.0002902	0.0017453
12	5.3132E−6	0.0015026	0.0001763	0.0015056	0.0006540	0.0001769	0.0015026
13	2.3057E−6	0.0013072	0.0001073	0.0013099	0.0005253	0.0001078	0.0013069
14	1.0000E−6	0.0011473	0.0000653	0.0011498	0.0004283	0.0000658	0.0011468
15	4.3357E−7	0.0010149	0.0000397	0.0010172	0.0003539	0.0000401	0.0010142
16	1.8787E−7	0.0009040	0.0000242	0.0009061	0.0002959	0.0000244	0.0009031
17	8.1375E−8	0.0008102	0.0000147	0.0008122	0.0002499	0.0000149	0.0008093
18	3.5230E−8	0.0007302	8.9686E−6	0.0007321	0.0002131	9.1270E−6	0.0007292
19	1.5246E−8	0.0006614	5.4592E−6	0.0006631	0.0001832	5.5729E−6	0.0006603
20	6.5952E−9	0.0006018	3.3230E−6	0.0006035	0.0001588	3.4034E−6	0.0006007

For the individual risk model, we have obtained a pdf of the aggregated risks, which corresponds to a beta distribution of the second kind. Then, we have considered the collective risk model based on dependence. We have studied some relevant models with Poisson, negative binomial and logarithmic distributions as primary distributions. For the collective Pareto–Poisson model, the pdf is a function of the Kummer confluent hypergeometric function, and in the Pareto–negative binomial is a function of the Gauss hypergeometric function.

Finally, using a data set based on one-year vehicle insurance policies taken out in 2004–2005 (Jong and Heller, 2008), we have concluded that our collective dependent models outperform other collective models considered in the actuarial literature in terms of AIC and CAIC statistics.

Acknowledgments

The authors thank the Ministerio de Economía y Competitividad (projects ECO2013-48326-C2-2-P, JMS, FP, VJ and ECO2013-47092 EGD) for partial support of this work. We are grateful for the constructive suggestions provided by the reviewers, which have improved the paper. The authors thank Prof. Carol Alexander for her comments and suggestions, which have improved the paper. In addition, this work is part of the Research Project APIE 1/2015-17 (JMS, FP, VJ): "New methods for the empirical analysis of financial markets" of the Santander Financial Institute (SANFI) of UCEIF Foundation resolved by the University of Cantabria and funded with sponsorship from Fundación Banco Santander.

Appendix

Proof of Proposition 1. Random variables $X_1, \dots, X_k$ are said to be associated if

$$\mathrm{cov}(\phi(X_1, \dots, X_k), \psi(X_1, \dots, X_k)) \geq 0$$

for all increasing functions ϕ, ψ for which the covariance exists. Now, the proof of Proposition 1 is based on the fact that the random variables $X_1, \dots, X_n$ are increasing functions of independent random variables (Esary et al., 1967).

Proof of Theorem 1. We have,

$$S_n = X_1 + \cdots + X_n = \beta \frac{Y_1 + \cdots + Y_n}{Y_\alpha},$$

and the distribution of the numerator is a $\mathcal{G}a(n)$ and the denominator is $\mathcal{G}a(\alpha)$, where both random variables are independent. Consequently, the ratio is a beta distribution of the second kind with pdf given by (6).

Proof of Theorem 2. The proof is direct,

$$\begin{aligned} F_{S_N}(x) &= \Pr(S_N \leq x) \\ &= \sum_{n=0}^{\infty} \Pr(S_N \leq x | N = n) \Pr(N = n) \\ &= \sum_{n=0}^{\infty} p_n F_X^{(n)}(x), \end{aligned}$$

where now $F_X^{(n)}(x)$ represents the cdf of the convolution of the n dependent claims $(X_1, \dots, X_n)$.

Proof of Lemma 2. The tail moments can be expressed as,

$$E(X^r | X > a) = \frac{\int_a^\infty x^r dF_X(x)}{1 - F_X(a)},$$

where $a = \mathrm{VaR}[X; u]$. Using the incomplete distribution of the second kind beta distribution we obtain the result.

Proof of Lemma 3. The formula for $E(S_N)$ is direct. Formula for $var(S_N)$ can be obtained using the identity $var(S_N) = E[var[S_N|N]] + var[E[S_N|N]]$.

Proof of Theorem 3. If $x = 0$, $f_{S_N}(0) = \Pr(N = 0) = e^{-\lambda}$. If $x > 0$ and calling $z = \frac{\lambda x/\beta}{1+x/\beta}$,

$$\begin{aligned} f_{S_N}(x) &= \sum_{n=1}^{\infty} f_X^{(n)}(x) \Pr(N = n) \\ &= \sum_{n=1}^{\infty} \frac{x^{n-1}}{\beta^n B(n, \alpha)(1 + x/\beta)^{n+\alpha}} \frac{e^{-\lambda}\lambda^n}{n!} \\ &= \frac{x^{-1} e^{-\lambda}}{\Gamma(\alpha)(1 + x/\beta)^\alpha} \sum_{n=1}^{\infty} \frac{\Gamma(n + \alpha)}{\Gamma(n) n!} \left(\frac{\lambda x/\beta}{1 + x/\beta} \right)^n \\ &= \frac{x^{-1} e^{-\lambda}}{\Gamma(\alpha)(1 + x/\beta)^\alpha} z \Gamma(\alpha + 1) {}_1F_1[1 + \alpha; 2; z], \end{aligned}$$

and we obtain (11).

Proof of Corollary 1. The proof is direct using Theorem 4 and taking into account that,

$$ {}_2F_1\left[2, 1+\alpha; 2; z\right] = (1-z)^{-\alpha-1}. $$

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: B.N. Petrov and F. Csaki (Eds.) 2nd International Symposium on Information Theory, pp. 267–281.

Albrecher, H., Teugels, J., 2006. Exponential behavior in the presence of dependence in risk theory. J. Appl. Probab. 43, 265–285.

Arnold, B.C., 1983. Pareto Distributions. International Cooperative Publishing House, Fairland, MD.

Arnold, B.C., 1987. Bivariate distributions with Pareto conditionals. Statist. Probab. Lett. 5, 263–266.

Arnold, B.C., 2015. Pareto Distributions, second ed. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Boca Ratón, FL.

Arnold, B.C., Castillo, E., Sarabia, J.M., 1993. Multivariate distributions with generalized Pareto conditionals. Statist. Probab. Lett. 17, 361–368.

Arnold, B.C., Castillo, E., Sarabia, J.M., 2001. Conditionally specified distributions: An introduction (with discussion). Statist. Sci. 16, 151–169.

Asimit, A.V., Furman, E., Vernic, R., 2010. On a multivariate Pareto distribution. Insurance Math. Econom. 46, 308–316.

Boudreault, M., Cossette, H., Landriault, D., Marceau, E., 2006. On a risk model with dependence between interclaim arrivals and claim sizes. Scand. Actuar. J. 301–323.

Bozdogan, H., 1987. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika 52, 345–370.

Chiragiev, A., Landsman, Z., 2009. Multivariate flexible Pareto model: Dependency structure, properties and characterizations. Statist. Probab. Lett. 79, 1733–1743.

Cossette, H., Coté, M.-P., Marceau, E., Moutanabbir, K., 2013. Multivariate distribution defined with Farlie–Gumbel–Morgenstern copula and mixed Erlang marginals: Aggregation and capital allocation. Insurance Math. Econom. 52, 560–572.

Cossette, H., Landriault, D., Marceau, E., 2004. Compound binomial risk model in a Markovian environment. Insurance Math. Econom. 35, 425–443.

Cossette, H., Marceau, E., Marri, F., 2008. On the compound Poisson risk model with dependence based on a generalized FarlieGumbelMorgenstern copula. Insurance Math. Econom. 43, 444–455.

Esary, J.D., Proschan, F., Walkup, D.W., 1967. Association of random variables, with applications. Ann. Math. Statist. 38, 1466–1474.

Genest, G., Marceau, E., Mesfioui, M., 2003. Compound Poisson approximations for individual models with dependent risks. Insurance Math. Econom. 32, 73–91.

Gerber, H.U., 1988. Mathematical fun with the compound binomial process. ASTIN Bull. 18, 161–168.

Gómez-Déniz, E., Calderín, E., 2014. Unconditional distributions obtained from conditional specifications models with applications in risk theory. Scand. Actuar. J. 2014 (7), 602–619.

Guillén, M., Sarabia, J.M., Prieto, F., 2013. Simple risk measure calculations for sums of positive random variables. Insurance Math. Econom. 53, 273–280.

Jong, P., Heller, G., 2008. Generalized Linear Models for Insurance Data. Cambridge University Press.

Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M., 2001. Modern Actuarial Risk Theory. Kluwer Academic Publishers, Boston.

Klugman, S.A., Panjer, H.H., Willmot, G.E., 2008. Loss Models. From Data to Decisions, third ed. John Wiley, New York.

Ramsay, C.M., 2006. The distribution of sums of certain i.i.d. Pareto variates. Comm. Statist. Theory Methods 35, 395–405.

Ramsay, C.M., 2009. The distribution of compound sums of Pareto distributed losses. Scand. Actuar. J. 2009, 27–37.

Rolski, T., Schmidli, H., Schmidt, V., Teugel, J., 1999. Stochastic Processes for Insurance and Finance. John Wiley and Sons, New York.

Sarabia, J.M., Guillén, M., 2008. Joint modelling of the total amount and the number of claims by conditionals. Insurance Math. Econom. 43, 466–473.

Shengwang, M., Wei, Y., Whitmore, G.A., 1999. Accounting for individual overdispersion in a bonus-malus automobile insurance system. ASTIN Bull. 29 (2), 327–337.

Wolfram, S., 2003. The Mathematica Book. Wolfram Media, Inc.