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
0167-6687/

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

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

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,

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

$$f(x_1,\dots,x_n;\alpha,n)=\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\beta^n}\frac{1}{\left(1+\sum_{i=1}^n x_i/\beta\right)^{\alpha+n}},\quad x_1,\dots,x_n>0. \tag{5}$$

  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 $\boldsymbol{X}=(X_1,\dots,X_n)^\top$ 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,

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

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

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

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