ON A MULTIVARIATE PARETO DISTRIBUTION [1]

Alexandru V. Asimit

School of Mathematics, University of Manchester, Manchester M13 9PL, United Kingdom. E-mail: Vali.Asimit@manchester.ac.uk

Edward Furman[2, 3]

Department of Mathematics and Statistics, York University, Toronto, Ontario M3J 1P3, Canada. E-mail: efurman@mathstat.yorku.ca

Raluca Vernic

Faculty of Mathematics and Informatics, Ovidius University of Constanta, Constanta, Romania. E-mail: rvernic@univ-ovidius.ro

Abstract. A multivariate distribution possessing arbitrarily parameterized Pareto margins is formulated and studied. The distribution is believed to allow for an adequate modeling of dependent heavy tailed risks with a non-zero probability of simultaneous loss. Numerous links to certain nowadays existing probabilistic models, as well as seemingly useful characteristic results are proved. Expressions for, e.g., decumulative distribution functions, densities, (joint) moments and regressions are developed. An application to the classical pricing problem is considered, and some formulas are derived using the recently introduced economic weighted premium calculation principles.

Keywords and phrases: Multivariate Pareto distributions, characterizations, mixtures, dependence, simultaneous loss, economic weighted pricing.

Mathematics Subject Classification: IM01, IM10, IM54.

---

[1]We are indebted to Professors Gheoghita Zbaganu (University of Bucharest, Romania) and Ričardas Zitikis (University of Western Ontario, Canada) for many illuminating discussions, and we also sincerely thank the anonymous referee for helpful comments and suggestions.

[2]The financial support of Natural Sciences and Engineering Research Council (NSERC) of Canada is gratefully appreciated.

[3]Corresponding author.

1. Introduction

Pareto probability laws were formulated by Pareto (1897) and initially dealt with the distribution of income over a population. They are nowadays popular in describing a variety of other phenomena such as city population sizes, occurrence of natural resources, stock price fluctuations, to name a few.

In actuarial mathematics, the excess-of-loss random variable (rv) $Y = X - d$ given that $X > d$, possesses the decumulative distribution function (ddf)

$$\overline{F}_{X,d}(x) = \mathbf{P}[X - d > x|\ X > d], \tag{1.1}$$

and it is of central importance in both life and non-life insurance contexts. Loosely speaking, ddf (1.1) on the one hand describes a risk inherent in a reinsurance treaty bearing a deductible d, and on the other hand it corresponds to the age-at-death of a new born child given survival to age d. Interestingly, it can be shown that for large d, appropriately chosen function $\sigma(d) > 0$ and a constant $\alpha > 0$, the approximation

$$\overline{F}_{X,d}(x) \approx \left(1 + \frac{x}{\sigma(d)}\right)^{-\alpha},\ x > 0 \tag{1.2}$$

holds for a large class of distribution functions (see, e.g., Balkema and de Haan, 1974; Pickands, 1975). Thus, the excess-of-loss rv can be generally approximated by a Pareto of the second kind. This, in turn, immediately hints at utilizing this probability law for modeling, e.g., catastrophe insurance treaties, which are known to be characterized by rather heavy tailed risks.

In view of the high popularity of the univariate Pareto distributions and the aforementioned implication of equation (1.2), it is quite natural to ask for a multivariate extension. Indeed, the concept of dependence, unfairly neglected in the 'classical' actuarial science, has been receiving its merited attention in the recent years, thus leading to numerous papers touching on various aspects of the multivariate distribution theory with applications to insurance (see, e.g., Vernic, 1997, 2000; Pfeifer and Nešlehová, 2004; Bauerle and Grubel, 2005; Roger et al., 2005; Centeno, 2005; Furman, 2008; Furman and Landsman, 2008, 2009; Chiragiev and Landsman, 2009).

Recently, probably due to the observation formulated by equation (1.2), the multivariate Pareto distribution of the second kind having the ddf

$$\overline{F}(x_1, \ldots, x_n) = \left(1 + \sum_{j=1}^{n} \frac{x_j - \mu_j}{\sigma_j}\right)^{-\alpha},\ x_j > \mu_j, \tag{1.3}$$

where $\mu_j \in (-\infty, \infty) := \mathbf{R}$, $\sigma_j \in (0, \infty) := \mathbf{R}_+$ and $j = 1, 2, \ldots, n$, has been employed to model insurance risks in the context of risk capital allocations and optimal reinsurance retentions by Chiragiev and Landsman (2007) (see, also, Vernic, 2009) and Cai and Tan (2007), respectively. In what follows, we denote the random vectors distributed multivariate Pareto of the second kind with ddf (1.3) as $\mathbf{X} \backsim Pa_n^*(II)(\boldsymbol{\mu}, \boldsymbol{\sigma}, \alpha)$, where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)' \in (-\infty, \infty)^n := \mathbf{R}^n$ and $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)' \in (0, \infty)^n := \mathbf{R}_+^n$ are two constant vectors, and α is a positive constant. Although the model seems to be quite tractable, it inconveniently results in common shape parameters for all univariate margins X_j, $j = 1, 2, \ldots, n$, and thus imposes the somewhat restrictive correlation coefficient given by

$$\mathbf{Corr}[X_j, X_l] = \frac{1}{\alpha}, \text{ for all } j \neq l \in \{1, 2, \ldots, n\} \text{ and } \alpha > 2.$$

Another well-known inconvenience associated with (1.3) is that it does not allow to model independent Pareto distributed risks.

In the present paper we propose an alternative to (1.3) multivariate generalization of the univariate Pareto probability laws. The probabilistic model introduced herein is constructed using the multivariate reduction method (see, e.g., Furman and Landsman, 2009; and references therein). The motivation for, as well as the interpretation of the aforementioned construction stem from, e.g., the background economy and the common shock models (see, e.g., Tsanakas, 2008; Boucher et al., 2008). The resulting multivariate Pareto, denoted in the sequel as $\mathbf{X} \backsim Pa_n(II)(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \alpha)$, with $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$ and α as before, and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)' \in \mathbf{R}_+^n$, is marginally closed, allows for non-negative probabilities of simultaneous losses and possesses a more flexible than (1.3) dependence structure. In addition, setting $\alpha \equiv 0$ yields a probabilistic model having independent Pareto distributed margins.

The rest of the paper is organized as follows: In Section 2 the multivariate Pareto of interest is introduced, and its various properties are derived. In Section 3 the discussion is specialized to the bivariate case, for which the (joint) moments, the conditional distributions along with the conditional moments are developed. Finally, in Section 4 it is shown that the regression function of the multivariate Pareto introduced herein is 'separable', which is then employed to facilitate its applications to insurance pricing. Section 5 concludes the paper.

2. The multivariate Pareto of the second kind

In the sequel, we fix the probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and we are interested in constructing a random vector $\mathbf{X} = (X_1, \ldots, X_n)'$, which is a map from the aformentioned space into the n-dimensional Borel space $(\mathbf{R}_+^n, \mathcal{B}^n)$, such that the j-th coordinate, $j = 1, \ldots, n$, of $\mathbf{X}$ is a univariate Pareto of the second kind. In other words, $X_j \backsim Pa(II)(\mu_j, \sigma_j, \alpha_j)$, and

$$\overline{F}_{X_j}(x) = \left(1 + \frac{x - \mu_j}{\sigma_j}\right)^{-\alpha_j}, \ x > \mu_j, \text{ and } \ \sigma_j, \ \alpha_j > 0.$$

To this end, let $\mathbf{Y} = (Y_0, Y_1, \ldots, Y_n)'$ be an $(n+1)$ variate random vector possessing mutually independent univariate Pareto margins $Y_0 \backsim Pa(II)(0, 1, \alpha_0)$ and $Y_j \backsim Pa(II)(\mu_j, \sigma_j, \alpha_j), \ j = 1, \ldots, n$. Denote by T the functional map $T : \mathbf{R}_+^{n+1} \to \mathbf{R}_+^n$.

Definition 2.1. *The random vector $\mathbf{X} = T(\mathbf{Y})$, such that*

$$X_j = \min\left(\sigma_j Y_0 + \mu_j, Y_j\right), \ j = 1, \ldots, n,$$

is said to follow the multivariate Pareto distribution of the second kind $Pa_n\,(II)\,(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \alpha_0)$, with $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, $\boldsymbol{\alpha}$ being constant vectors in $\mathbf{R}^n$ and $\mathbf{R}_+^n$, respectively, and $\alpha_0 > 0$.

Note 2.1. *It is possible to formulate the definition above for a slightly more general random vector $\mathbf{Y}$ by assuming $Y_0 \backsim Pa(II)(\mu_0, \sigma_0, \alpha_0)$, which is a non-standardized Pareto of the second kind. In such a case, the map $T : \mathbf{R}_+^{n+1} \to \mathbf{R}_+^n$ is such that*

$$X_j = \min\left(\frac{\sigma_j}{\sigma_0}\left(Y_0 - \mu_0\right) + \mu_j, Y_j\right), \quad j = 1, \ldots, n.$$

However, as it is readily observed from Proposition 2.1, neither μ_0 nor σ_0 influences the ddf of $\mathbf{X}$, and they are hence set to $\mu_0 \equiv 0$ and $\sigma_0 \equiv 1$ without loss of generality.

We note in passing that various particular cases of the just defined multivariate Pareto have been studied. Namely, when $\mu_j = \sigma_j \equiv const$ for $j = 1, \ldots, n$, the model is called the Marshall-Olkin type (see, Marshall and Olkin, 1967) multivariate Pareto in Kotz et al. (2000), and it has been explored by Muliere and Scarsini (1987) in the two dimensional, and by Hanagal (1996) in the multidimensional contexts.

We further derive the multivariate ddf of the distribution defined in Definition 2.1. To enhance the readability of the paper, we relegate the proof of the next proposition, as well as other technical proofs to the Appendix.

Proposition 2.1. *Let* $\mathbf{X} \backsim Pa_n(II)(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \alpha_0)$*. The ddf of* $\mathbf{X}$ *is then*

$$\bar{F}_{\mathbf{X}}(\mathbf{x}) = \left(1 + \max_{j=1,\ldots,n} \frac{x_j - \mu_j}{\sigma_j}\right)^{-\alpha_0} \prod_{j=1}^{n} \left(1 + \frac{x_j - \mu_j}{\sigma_j}\right)^{-\alpha_j}, x_j > \mu_j,\ j = 1, \ldots, n, \quad (2.1)$$

where μ_j *are real, and* $\sigma_j,\ \alpha_j,\ \alpha_0$ *are positive constants.*

Note 2.2. *It is straightforwardly observed that ddf (2.1) is invariant under linear transformations, and it thus belongs to a (multivariate) location-scale family of distributions. Consequently, in the sequel, we shall frequently use the so-called standardized form of the ddf, which certainly does not reduce the level of generality, but strongly contributes to the clarity of the results. More specifically, we shall often consider the multivariate Pareto* $\mathbf{Z} \backsim Pa_n(II)(\mathbf{0}, \mathbf{1}, \boldsymbol{\alpha}, \alpha_0)$*, with* $\boldsymbol{\mu} = \mathbf{0}$ *and* $\boldsymbol{\sigma} = \mathbf{1}$ *being* n*-variate constant vectors of zeros and ones, respectively, and the ddf*

$$\bar{F}_{\mathbf{Z}}(z_1, \ldots, z_n) = \left(1 + \max_{j=1,\ldots,n} z_j\right)^{-\alpha_0} \prod_{j=1}^{n} (1 + z_j)^{-\alpha_j}, z_j > 0,\ j = 1, \ldots, n. \quad (2.2)$$

It seems worth noticing that there exist alternative to Definition 2.1 ways to arrive at ddf (2.1). We formulate our observations in this context as the following propositions, and we leave their proofs to the reader.

Proposition 2.2. *Let* $\mathbf{Y} = (Y_1, \ldots, Y_n)' \backsim Exp_n(\boldsymbol{\alpha}, \alpha_0)$ *denote the* n *variate Marshall and Olkin exponential distribution with* $\boldsymbol{\alpha} \in \mathbf{R}_+^n$*,* $\alpha_0 \in \mathbf{R}_+$ *and the ddf*

$$\overline{F}_{\mathbf{Y}}(y_1, \ldots, y_n) = \exp\{-\alpha_1 y_1 - \cdots - \alpha_n y_n - \alpha_0 \max(y_1, \ldots, y_n)\}.$$

Then the n *variate random vector* $\mathbf{X}$ *with its* j*-th coordinate given by* $X_j = \mu_j + \sigma_j(e^{Y_j} - 1),\ j = 1, \ldots, n$ *has ddf (2.1).*

We have already emphasized that the multivariate Pareto of Hanagal (1996) is a special case of the one given in Definition 2.1. More specifically, to obtain the former distribution, we set $\boldsymbol{\mu} = \boldsymbol{\sigma} = (\sigma, \ldots, \sigma)' \in \mathbf{R}_+^n$ in the framework of the latter one.

Proposition 2.3. *Let* $\mathbf{Y} \backsim Pa_n(II)(\boldsymbol{\sigma}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \alpha_0)$ *denote the* n *variate Pareto of Hanagal (1996) with* $\boldsymbol{\sigma} = (\sigma, \ldots, \sigma)' \in \mathbf{R}_+^n$*,* $\boldsymbol{\alpha} \in \mathbf{R}_+^n$*,* $\alpha_0 \in \mathbf{R}_+$ *and the ddf*

$$\overline{F}_{\mathbf{Y}}(y_1, \ldots, y_n) = \left(\max_{j=1,\ldots,n} \frac{y_j}{\sigma}\right)^{-\alpha_0} \prod_{j=1}^{n} \left(\frac{y_j}{\sigma}\right)^{-\alpha_j}.$$

Then the n *variate random vector* $\mathbf{X}$ *with its* j*-th coordinate given by* $X_j = \mu_j + \sigma_j(Y_j/\sigma - 1)$ *has ddf (2.1).*

Our next observation touches on a mixture representation of the multivariate Pareto discussed herein. The counterpart of Proposition 2.4 is well documented for the multivariate Pareto of the second kind given by (1.3), and it has been extensively employed in actuarial science. We believe that the mixture representation below is new and worth noticing. Following Note 2.2, we formulate and prove the proposition for the standardized multivariate Pareto, i.e., for $\mathbf{Z} \backsim Pa_n(\mathbf{0}, \mathbf{1}, \boldsymbol{\alpha}, \alpha_0)$. The proof is sketched in the Appendix.

Proposition 2.4. *Let $\mathbf{Z}|(\boldsymbol{\Lambda}, \Lambda_0)' \backsim Exp_n(\boldsymbol{\Lambda}, \Lambda_0)$ denote, as before, the n variate Marshall and Olkin exponential distribution, and let $(\boldsymbol{\Lambda}, \Lambda_0)'$ be an $(n+1)$ dimensional random vector with independent univariate standardized gamma distributed coordinates $\Lambda_i \backsim Ga(\alpha_i, 1)$, $i = 0, 1, \ldots, n$. Then the unconditional distribution of $\mathbf{Z}$ is $Pa_n\,(II)\,(\mathbf{0}, \mathbf{1}, \boldsymbol{\alpha}, \alpha_0)$.*

It turns out that the multivariate Pareto possessing ddf (1.3) is to an extent a particular case of the one given by Definition 2.1. We formulate this observation as the following proposition. The proof in essence replicates the one of Proposition 2.4 and is thus omitted. Also, in a manner entirely similar to Note 2.2, we let $Pa_n^*\,(II)\,(\mathbf{0}, \mathbf{1}, \alpha)$ be the standardized counterpart of multivariate Pareto (1.3).

Proposition 2.5. *Let $\mathbf{Z}|\boldsymbol{\Lambda} \backsim Exp_n(\boldsymbol{\Lambda}, 0)$ be the n variate Marshall and Olkin exponential distribution, and let $\boldsymbol{\Lambda}$ be an n dimensional random vector possessing standardized gamma coordinates $\Lambda \backsim Ga(\alpha, 1)$. Then the unconditional distribution of $\mathbf{Z}$ is $Pa_n^*\,(II)\,(\mathbf{0}, \mathbf{1}, \alpha)$.*

Interestingly, it is possible to extend Proposition 2.4, and obtain a characteristic property of the multivariate Pareto distribution introduced in Definition 2.1.

Proposition 2.6. *If $\mathbf{Z} \backsim Pa_n\,(II)\,(\mathbf{0}, \mathbf{1}, \boldsymbol{\alpha}, \alpha_0)$, then there exist a random vector $\boldsymbol{\Xi}$ following the n variate Marshall and Olkin exponential distribution $Exp_n(\lambda_1, \ldots, \lambda_n, \lambda_0)$ and an $(n+1)$ variate random vector $\boldsymbol{\Lambda}$ possessing independent standardized gamma margins $Ga(\alpha_i, 1)$, $i = 0, 1, \ldots, n$, such that $\mathbf{X} = \boldsymbol{\Xi}|\boldsymbol{\Lambda}$.*

Propositions 2.4 and 2.6 establish an ‘if and only if’ result for the mixture representation of the multivariate Pareto of interest.

We next show that the introduced multivariate Pareto is marginally closed, which seems to be one of the basic requirements for a reasonable multivariate distribution aiming at (insurance) applications.

Proposition 2.7. *Let $\mathbf{X} \backsim Pa_n(II)(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \alpha_0)$. Then the j-th univariate margin is distributed Pareto of the second kind. Namely, $X_j \backsim Pa(II)\,(\mu_j, \sigma_j, \alpha_0 + \alpha_j)\,,\ j = 1, \ldots, n$.*

We note in passing that in a similar fashion, multivariate margins of any order follow multivariate Pareto distributions of the second kind with arbitrarily parameterized univariate margins.

One of the peculiarities associated with the multivariate Pareto introduced in Definition 2.1 is that, unlike many other multivariate distributions (including (1.3)), it allows for modeling simultaneously occurring losses. In the context of (1.3), we certainly have that $\mathbf{P}[X_1 = X_2 = \cdots = X_n] = 0$. More precisely speaking, this means that the multivariate Pareto introduced in Definition 2.1 is not absolutely continuous with respect to the Lebesgue measure on $\mathbf{R}_+^n$, and it thus has both an absolutely continuous and a singular part.

Distributions with singularities are rarely popular in the univariate case, however they can be easily motivated in the multivariate one. Indeed, specializing our discussion to the insurance context of interest and bearing in mind the n variate random vector $\mathbf{X} = (X_1, X_2, \ldots, X_n)'$, the event $\{X_1 = X_2 = \cdots = X_n\}$ can represent, e.g., the possibility of a simultaneous claim initiated by n business lines. Needless to say, the aforementioned event is also useful to plainly model the occurrence of equal losses, which is certainly impossible to implement in the framework of absolutely continuous multivariate distributions.

In the next proposition we quantitatively study the singularity phenomenon.

Proposition 2.8. *For $k \leq n$ and distinct $j_1, ..., j_k \in \{1, 2, \ldots, n\}$, we have that*

$$\mathbf{P}\left[\frac{X_{j_1} - \mu_{j_1}}{\sigma_{j_1}} = \cdots = \frac{X_{j_k} - \mu_{j_k}}{\sigma_{j_k}}\right] = \frac{\alpha_0}{\alpha_0 + \sum_{i=1}^k \alpha_{j_i}}.$$

We have already noticed that the ddf of $\mathbf{X} \backsim Pa_n(II)(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \alpha_0)$ is a mixture of the absolutely continuous and singular components. In view of this, the corresponding multivariate probability density function (pdf) can be derived by taking the appropriate derivative of the continuous component and utilizing Proposition 2.8.

Note 2.3. *The singular component is 0 if and only if $\alpha_0 \equiv 0$. In addition, the greater the dependence among risks is, the more substantial is the portion of the singularity.*

We further specialize our discussion to the bivariate case, where a decomposition of the decumulative distribution function into absolutely continuous and singular components can be conveniently elaborated.

3. The bivariate Pareto distribution of the second kind

We are henceforth interested in the bivariate counterpart of Definition 2.1. Clearly, in such a case ddf (2.1) reduces to

$$\bar{F}_{\mathbf{X}}(x_1,x_2)=\left(1+\max_{j=1,2}\frac{x_j-\mu_j}{\sigma_j}\right)^{-\alpha_0}\prod_{j=1}^{2}\left(1+\frac{x_j-\mu_j}{\sigma_j}\right)^{-\alpha_j},\ x_j>\mu_j,\ j=1,2, \tag{3.1}$$

or, in other words, to

$$\bar{F}_{\mathbf{X}}\left(x_1,x_n\right)=\begin{cases}\left(1+\dfrac{x_1-\mu_1}{\sigma_1}\right)^{-\alpha_{01}}\left(1+\dfrac{x_2-\mu_2}{\sigma_2}\right)^{-\alpha_2}, & \dfrac{x_1-\mu_1}{\sigma_1}\geq\dfrac{x_2-\mu_2}{\sigma_2}>0\\ \left(1+\dfrac{x_1-\mu_1}{\sigma_1}\right)^{-\alpha_1}\left(1+\dfrac{x_2-\mu_2}{\sigma_2}\right)^{-\alpha_{02}}, & \dfrac{x_2-\mu_2}{\sigma_2}\geq\dfrac{x_1-\mu_1}{\sigma_1}>0\end{cases},$$

where $\alpha_{0j}=\alpha_0+\alpha_j,\ j=1,\ 2.$

Note 3.1. *When formulated more accurately, bivariate ddf (3.1) is given by*

$$\bar{F}_{\mathbf{X}}(x_1,x_2)=\begin{cases}\left(1+\max_{j=1,2}\frac{x_j-\mu_j}{\sigma_j}\right)^{-\alpha_0}\prod_{j=1}^{2}\left(1+\frac{x_j-\mu_j}{\sigma_j}\right)^{-\alpha_j}, & x_j>\mu_j,\\ \left(1+\dfrac{x_1-\mu_1}{\sigma_1}\right)^{-\alpha_{01}}, & x_1>\mu_1,\ x_2\leq\mu_2\\ \left(1+\dfrac{x_2-\mu_2}{\sigma_2}\right)^{-\alpha_{02}}, & x_1\leq\mu_1,\ x_2>\mu_2\\ 1, & x_1\leq\mu_1,\ x_2\leq\mu_2\end{cases}.$$

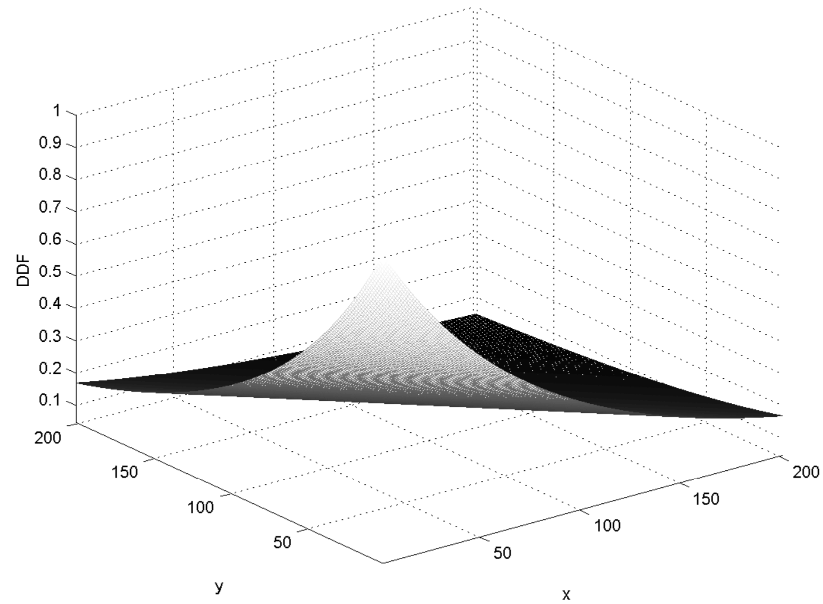
In what follows, we shall however consider the sub-domain $x_j>\mu_j,\ j=1,\ 2,$ only.

In Figure 1 we have depicted ddf (3.1) versus the bivariate counterpart of (1.3). The parameters of the two distributions are chosen in such a way that their means (3.3) are equal, as well as their variances (3.4). Also, the two models share same correlation value. Figure 1 conveniently visualizes the fact that the Pareto distribution introduced in this paper is not differentiable everywhere. In what follows, we elaborate on this phenomenon.

3.1. The decumulative distribution function decomposition.

Bivariate ddf (3.1) is not absolutely continuous with respect to the Lebesgue measure on $\mathbf{R}_+^2$. Namely, it consists of both the absolutely continuous and singular components with the latter concentrated on $\left\{(x_1,x_2):\dfrac{x_1-\mu_1}{\sigma_1}=\dfrac{x_2-\mu_2}{\sigma_2}\right\}$. Hence, the ddf can be decomposed as

$$\bar{F}_{\mathbf{X}}\left(\mathbf{x}\right)=a\bar{F}_{ac}\left(\mathbf{x}\right)+\left(1-a\right)\bar{F}_s\left(\mathbf{x}\right),\ 0\leq a\leq 1, \tag{3.2}$$



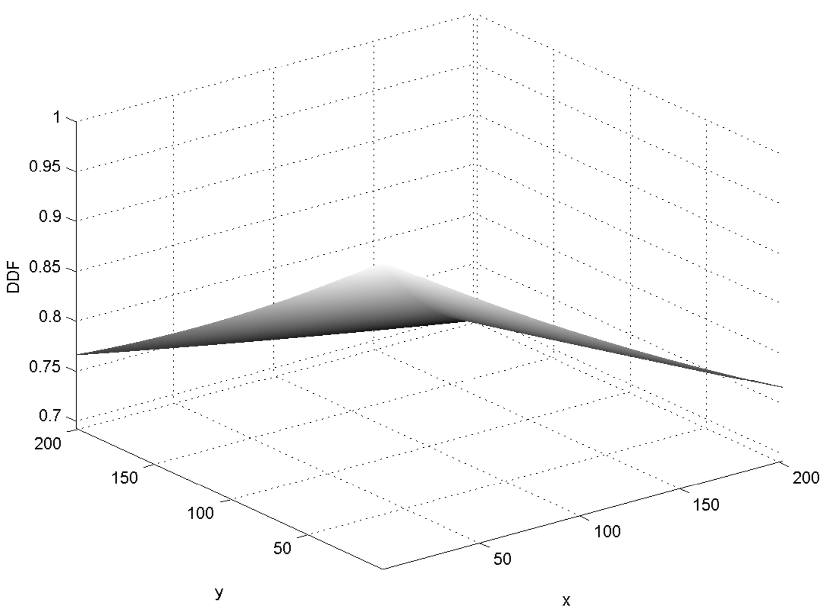


FIGURE 1. The ddf's $\overline{F}(x,y)$ of the classical Pareto (top pannel) and the new one (bottom panel).

where $\bar{F}_{ac}$ denotes the absolutely continuous component, and $\bar{F}_s$ stands for the singular one. The next proposition formally describes the phenomenon. The proof is given in the Appendix. From now on, we often use the following auxiliary notation $\alpha_+ = \alpha_0 + \alpha_1 + \alpha_2$.

Proposition 3.1. *Let* $\mathbf{X} \backsim Pa_2\,(II)\,(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \alpha_0)$*. Then the ddf of* $\mathbf{X}$ *is a mixture*

$$\bar{F}_{\mathbf{X}}\,(\mathbf{x}) = \frac{\alpha_1 + \alpha_2}{\alpha_+} \bar{F}_{ac}\,(\mathbf{x}) + \frac{\alpha_0}{\alpha_+} \bar{F}_s\,(\mathbf{x})\,,$$

where the singular component $\bar{F}_s(\mathbf{x})$ *concentrates its mass on the line* $(x_1-\mu_1)/\sigma_1 = (x_2-\mu_2)/\sigma_2$, *and is given by*

$$\bar{F}_s(\mathbf{x}) = \left(1+\max_{i=1,2}\frac{x_i-\mu_i}{\sigma_i}\right)^{-\alpha_+}, x_j > \mu_j,\ j=1,2,$$

while the absolutely continuous component is formulated, for $x_j > \mu_j,\ j=1,2,$ *as*

$$\bar{F}_{ac}(\mathbf{x}) = \frac{\alpha_+}{\alpha_1+\alpha_2}\left(1+\max_{i=1,2}\frac{x_i-\mu_i}{\sigma_i}\right)^{-\alpha_0}\prod_{i=1}^{2}\left(1+\frac{x_i-\mu_i}{\sigma_i}\right)^{-\alpha_i} - \frac{\alpha_0}{\alpha_1+\alpha_2}\left(1+\max_{i=1,2}\frac{x_i-\mu_i}{\sigma_i}\right)^{-\alpha_+}.$$

Note 3.2. *The pdf of the absolutely continuous component is obtained by straightforward differentiation, i.e.,*

$$\frac{\partial^2 \bar{F}_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \partial x_2} = \frac{\alpha_1+\alpha_2}{\alpha_+} f_{ac}(\mathbf{x})$$

$$= \begin{cases} \frac{\alpha_{01}\alpha_2}{\sigma_1\sigma_2}\left(1+\frac{x_1-\mu_1}{\sigma_1}\right)^{-(\alpha_{01}+1)}\left(1+\frac{x_2-\mu_2}{\sigma_2}\right)^{-(\alpha_2+1)}, & \frac{x_1-\mu_1}{\sigma_1} > \frac{x_2-\mu_2}{\sigma_2} > 0 \\ \frac{\alpha_{02}\alpha_1}{\sigma_1\sigma_2}\left(1+\frac{x_1-\mu_1}{\sigma_1}\right)^{-(\alpha_1+1)}\left(1+\frac{x_2-\mu_2}{\sigma_2}\right)^{-(\alpha_{02}+1)}, & \frac{x_2-\mu_2}{\sigma_2} > \frac{x_1-\mu_1}{\sigma_1} > 0 \end{cases}.$$

The pdf is employed in Propositions 3.2 and 3.3 below.

3.2. **Moments.** The marginal moments of $\mathbf{X} \backsim Pa_2(II)(\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\alpha},\alpha_0)$ are readily obtained, for $j=1,\ 2$, as

$$\mathbf{E}[X_j] = \mu_j + \frac{\sigma_j}{\alpha_{0j}-1},\quad \text{for } \alpha_{0j} > 1 \tag{3.3}$$

and

$$\mathbf{Var}[X_j] = \frac{\alpha_{0j}\sigma_j^2}{(\alpha_{0j}-1)^2(\alpha_{0j}-2)},\quad \text{for } \alpha_{0j} > 2. \tag{3.4}$$

The covariance of say X_1 and X_2 can also be derived, as we show with some effort in the next proposition.

Proposition 3.2. *Let* $\mathbf{X} \backsim Pa_2(II)(\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\alpha},\alpha_0)$, *and assume that* $\alpha_{0j} > 1,\ j=1,\ 2$ *and* $\alpha_+ > 2$. *Then*

$$\mathbf{Cov}[X_1,X_2] = \frac{\alpha_0\sigma_1\sigma_2}{(\alpha_{01}-1)(\alpha_{02}-1)(\alpha_+-2)}.$$

Moreover, if $\alpha_{0j} > 2,\ j=1,\ 2$, *then the correlation coefficient is*

$$\mathbf{Corr}[X_1,X_2] = \frac{\alpha_0}{\alpha_+-2}\sqrt{\frac{(\alpha_{01}-2)(\alpha_{02}-2)}{\alpha_{01}\alpha_{02}}}.$$

Note 3.3. *As a consequence of Proposition 3.2, we readily have that for fixed* α_1 *and* α_2 *and letting* $\alpha_0 \to \infty$, *the correlation between* X_1 *and* X_2 *varies from* 0 *to* 1.

3.3. **Conditional distributions and expectations.** Let $\mathbf{X} \backsim Pa_2(II)(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \alpha_0)$. Then the conditional distribution of X_1 given $X_2 = x_2$ has an absolutely continuous component and a discrete one. The following proposition formulates its ddf. The proof is in the Appendix.

Proposition 3.3. *The conditional ddf of X_1 given $X_2 = x_2$, $x_2 > \mu_2$, is*

$$\bar{F}_{X_1|X_2}\left(x_1 \left| x_2\right.\right) = \begin{cases} \dfrac{\alpha_2}{\alpha_{02}}\left(1+\dfrac{x_2-\mu_2}{\sigma_2}\right)^{\alpha_0}\left(1+\dfrac{x_1-\mu_1}{\sigma_1}\right)^{-\alpha_{01}}, & \dfrac{x_1-\mu_1}{\sigma_1} > \dfrac{x_2-\mu_2}{\sigma_2} \\ \left(1+\dfrac{x_1-\mu_1}{\sigma_1}\right)^{-\alpha_1}, & 0 < \dfrac{x_1-\mu_1}{\sigma_1} \le \dfrac{x_2-\mu_2}{\sigma_2} \end{cases}.$$

We then immediately obtain the next corollary which yields the conditional expectation.

Corollary 3.1. *For $\alpha_{01} > 1$, the regression of X_1 on X_2 is given by*

$$\mathbf{E}\left[X_1 \left| X_2 = x_2\right.\right] = \begin{cases} \mu_1 + \dfrac{\sigma_1}{\alpha_1 - 1}\left(1+\dfrac{\alpha_0\left(1-\alpha_+\right)}{\alpha_{02}\left(\alpha_{01}-1\right)}\left(1+\dfrac{x_2-\mu_2}{\sigma_2}\right)^{-(\alpha_1-1)}\right), & \text{if } \alpha_1 \neq 1 \\ \mu_1 + \sigma_1\left(\ln\left(1+\dfrac{x_2-\mu_2}{\sigma_2}\right)+\dfrac{\alpha_2}{\alpha_0\alpha_{02}}\right), & \text{if } \alpha_1 = 1 \end{cases},$$

where $x_2 > \mu_2$.

In the next section we relate Corollary 3.1 to certain issues of (insurance) pricing.

4. Applications to (insurance) pricing

Let $r_{X_1|X_2}(x_2) = \mathbf{E}[X_1 - \mathbf{E}[X_1]|X_2 = x_2]$ denote the centered regression function of X_1 on X_2. The following definition is a simplified (one dimensional) version of Definition 2.1 of Furman and Zitikis (2009).

Definition 4.1. *The centered regression function $r_{X_1|X_2}$ is called separable if it admits the decomposition*

$$r_{X_1|X_2}(x_2) = C(F_{(X_1,X_2)})q(x_2, F_{X_1}, F_{X_2}) \tag{4.1}$$

for F_{X_2}-almost all $x_2 \in \mathbf{R}$, where $C(F_{(X_1,X_2)})$ is a constant which depends on the joint cdf $F_{(X_1,X_2)}$, and $x_2 \mapsto q(x_2, F_{X_1}, F_{X_2}) \in \mathbf{R}$ is a function that may depend only on the marginal cdf's F_{X_1} and F_{X_2}.

It has turned out that separable regression functions considerably facilitate the analytic derivations of a great variety of pricing functionals. Following Corollary 3.1, we readily observe that the centered regression function of the multivariate Pareto of interest is separable. Namely, for e.g., $(Z_1, Z_2) \backsim Pa_2\left(II\right)\left(\mathbf{0}, \mathbf{1}, \boldsymbol{\alpha}, \alpha_0\right)$, $\alpha_{01} > 1$, $\alpha_{02} > 2$, $\alpha_1 \neq 1$, we have that

$$C(F_{(Z_1,Z_2)}) = \frac{\mathbf{Cov}[Z_1, Z_2]}{\mathbf{Var}[Z_2]} \frac{(\alpha_+ - 2)(\alpha_+ - 1)}{(\alpha_1 - 1)(\alpha_{02} - 2)(\alpha_{02} - 1)}, \tag{4.2}$$

and

$$q(x_2, F_{Z_1}, F_{Z_2}) = \left(\frac{\alpha_{02}}{\alpha_+ - 1} - (1 + x_2)^{1-\alpha_1} \right). \tag{4.3}$$

Similarly, the regression function of the non-standardized bivariate Pareto is also separable.

In Figure 2, we have depicted the regression functions of the classical bivariate Pareto versus the one of the model introduced in this paper. Unlike the former, the latter multivariate Pareto has a non-linear regression. We note in passing, that the parameters of the two distributions are chosen in such a way that their means are equal, as well as their variances. The models share same correlation value.

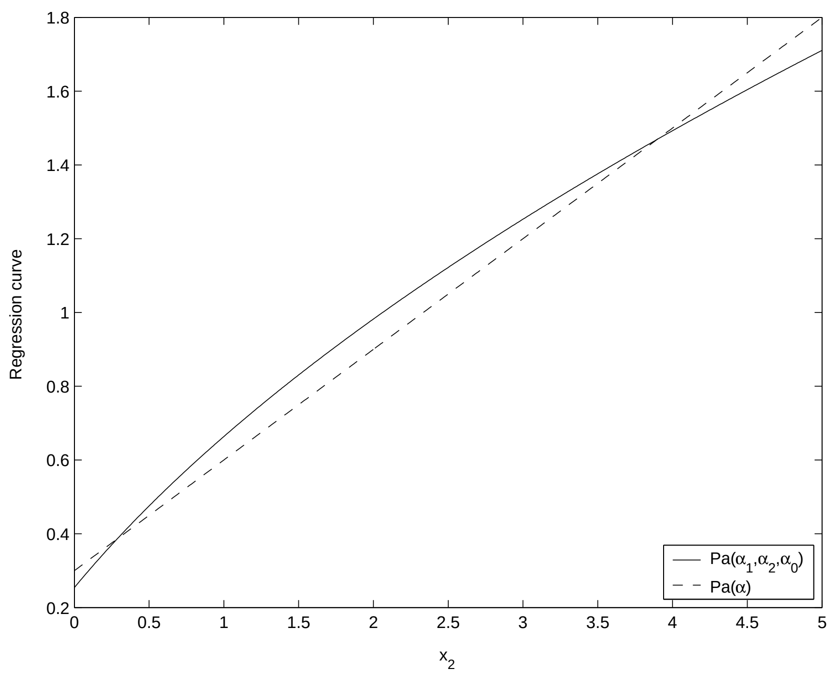


FIGURE 2. Regression curves of the classical and new bivariate Pareto distributions.

4.1. **Insurance pricing.** Let $\mathcal{X}$ denote the set of (insurance) risks. Then following Furman and Zitikis (2008a), we recall that for $X \in \mathcal{X}$ and a non-decreasing weight function $w : \mathbf{R} \to \mathbf{R}_+$, the actuarial weighted premium calculation principle is formulated as the map $\pi_w : \mathcal{X} \to [0, \infty]$, such that

$$\pi_w[X] = \frac{\mathbf{E}[Xw(X)]}{\mathbf{E}[w(X)]}, \tag{4.4}$$

where we assume that $0 < \mathbf{E}[w(X)] < \infty$. Functional (4.4) unifies such popular premium calculation principles as the Value-at-Risk, conditional tail expectation, Esscher, Kamps' and under mild conditions the distorted premiums, and it thus allows for a common treatment of all of the above.

It should be noted, that functional (4.4) establishes a classical actuarial pricing principle, since it depends on X only. Departing from this rather restrictive requirement, Furman and Zitikis (2008b) introduced and studied the economic weighted pricing functional $\Pi_w : \mathcal{X} \times \mathcal{X} \to [0, \infty]$ formulated as

$$\Pi_w[X_1, X_2] = \frac{\mathbf{E}[X_1 w(X_2)]}{\mathbf{E}[w(X_2)]}, \tag{4.5}$$

for $0 < \mathbf{E}[w(X_2)] < \infty$, and X_1 and X_2 in $\mathcal{X}$. They also showed that, for $D_w[F_{X_1}, F_{X_2}] = \mathbf{E}[q(X_2, F_{X_1}, F_{X_2}) w(X_2)]$, the following relation between Π_w and π_w holds

$$\Pi_w[X_1, X_2] \quad = \quad \mathbf{E}[X_1] + \frac{C(F_{(X_1,X_2)}) D_w(F_{X_1}, F_{X_2})}{\mathbf{Cov}[X_2, w(X_2)]} \left(\pi_w[X_2] - \mathbf{E}[X_2]\right). \tag{4.6}$$

Consequently, to calculate the value of $\Pi_w[X_1, X_2]$ in the case of the classical multivariate Pareto, we use the fact that it has linear centered regression function of the form (see, e.g., Arnold, 1983),

$$r^*_{X_1|X_2}(x_2) = \frac{\mathbf{Cov}[X_1, X_2]}{\mathbf{Var}[X_2]}(x_2 - \mathbf{E}[X_2]),$$

which is of course separable with $C(F_{(X_1,X_2)}) = \mathbf{Cov}[X_1, X_2]/\mathbf{Var}[X_2]$, $q(x_2, F_{X_1}, F_{X_2}) = x_2 - \mathbf{E}[X_2]$, and thus yields $D_w[F_{X_1}, F_{X_2}] = \mathbf{Cov}[X_2, w(X_2)]$. Consequently, the value of the economic pricing functional readily follows for an arbitrary choice of w using equation (4.6).

In a similar manner, we calculate the value of the economic pricing functional for the multivariate Pareto distribution introduced in this paper. More specifically, we note that in this case $C(F_{(X_1,X_2)})$ and $D_w[F_{X_1}, F_{X_2}]$ follow from equations (4.2) and (4.3), or more generally from Corollary 3.1. Hence, the value of the economic pricing functional once again readily follows for an arbitrary choice of w using equation (4.6).

As it is seen from the discussion above, the task of evaluating the economic weighted pricing functional is quite straightforward in the context of both the classical multivariate Pareto and the new one having the non-linear regression function. To illustrate the procedure numerically, we specialize our discussion to the conditional tail expectation pricing principle, which is the (economic) weighted pricing functional with $w(x) = \mathbf{1}\{x > VaR_p[X]\}$, where $VaR_p[X] = \inf\{x : F(x) \geq p\}$ denotes the well-known Value-at-Risk

risk measure. In this respect and given a random pair (X_1, X_2), we are interested in calculating the price of X_1.

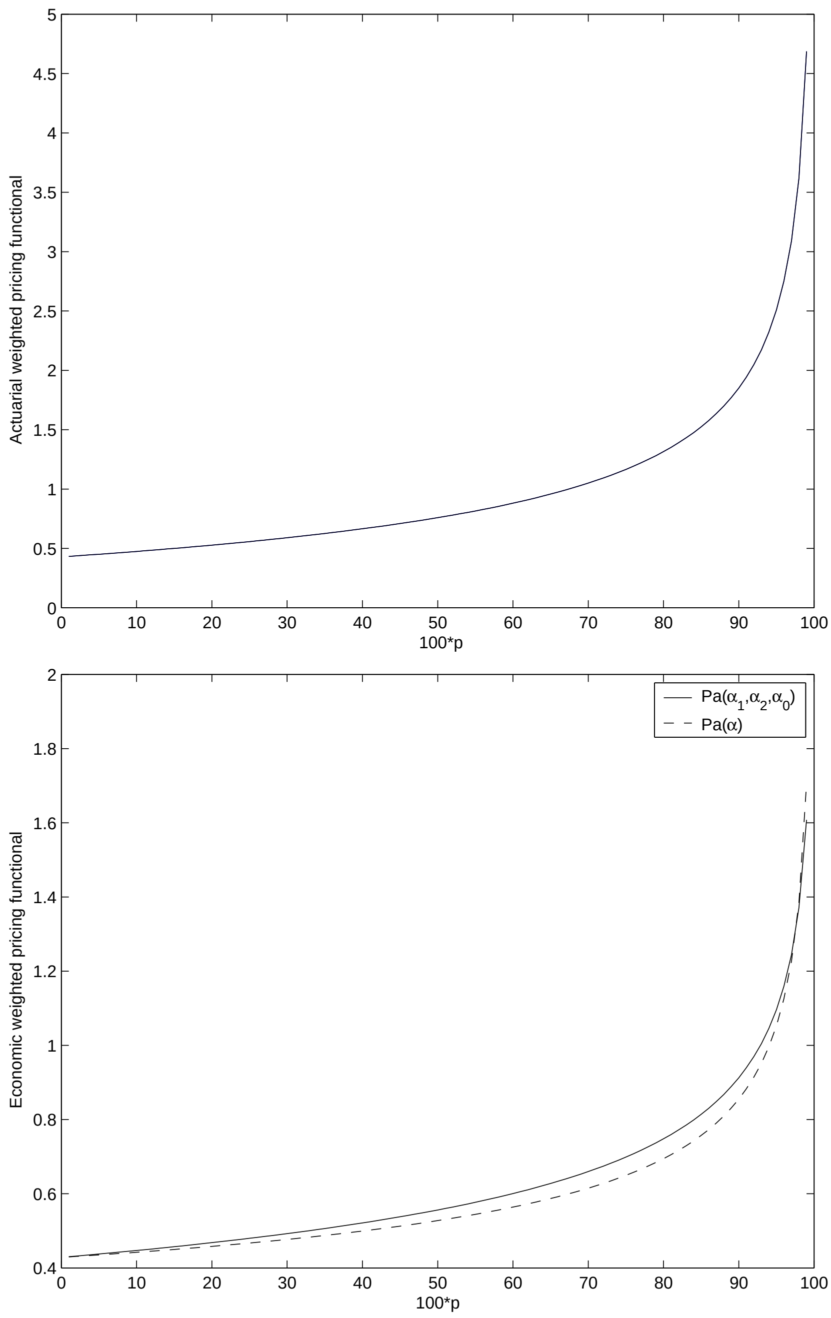


FIGURE 3. The actuarial (top panel) and the economic (bottom panel) weighted pricing functionals for the first coordinate of (X_1, X_2).

We have depicted the results in Figure 3. The top panel shows the values of the actuarial pricing functional $CTE_p[X_1] = \mathbf{E}[X_1 | X_1 > VaR_p[X_1]]$ for various values of p. Actuarial pricing functionals do not make any use of the dependence of X_1 on X_2, therefore they coincide for both Pareto's.

The importance of the dependence structure is however readily observed from the bottom panel of Figure 3, where the economic pricing functionals $CTE_p[X_1, X_2] = \mathbf{E}[X_1|X_2 > VaR_p[X_2]]$ are visualized for the classical multivariate Pareto and the new one introduced in this paper. To make the problem more interesting, we set $\mathbf{E}[X_1] = \mathbf{E}[X_2]$ and also $\mathbf{Var}[X_1] = \mathbf{Var}[X_2]$ for both models, that in addition share same correlation value. Nevertheless, the implications of the dependence structures are evident. Namely, it is clear that in this case, mistakenly assuming linear dependence between X_1 and X_2 can result in underpricing of X_1 for moderate values of p and overpricing of X_1 for the so-called tail values of p.

Finally, the actuarial CTE pricing functional is well-known to be at least as high as its economic counterpart, i.e., the bound $CTE_q[X_1, X_2] \leq CTE_q[X_1]$ holds for any random pair (X_1, X_2) and level of confidence q independently on the joint cdf $F_{(X_1,X_2)}$ (see, Furman and Zitikis, 2008b). The bound is reflected in Figure 4, which combines both panels of Figure 3.

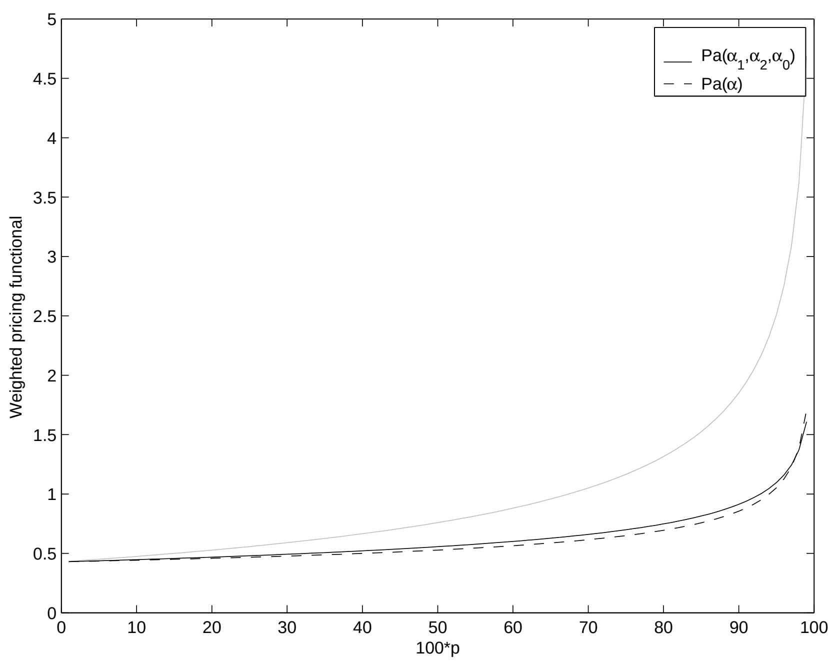


FIGURE 4. The actuarial (gray) and the economic (black) CTE pricing functionals for the first coordinate of (X_1, X_2).

5. CONCLUSIONS

Generally speaking, there exist a great variety of methods of constructing multivariate dependent probabilistic structures possessing specific univariate margins of interest, with the copula-based approaches being arguably the most popular in nowadays applied actuarial science and practice (see, e.g., Frees and Valdez, 1998; and Cherubini et al., 2004). Indeed, in many situations copulas can be quite efficient in the way they separate the dependence structure and the margins, thus supplying an attractive route for modeling dependent risks.

However, there is always a trade-off between the approximation level provided by the model and its analytic complexity. In this respect, copula-based multivariate probabilistic structures frequently suffer from analytic intractability (see, e.g., Song, 2000).

In this paper a multivariate distribution possessing arbitrarily parameterized Pareto margins has been formulated and studied. The distribution is unimodal and positively skewed, and it conveniently allows for modeling the probability of simultaneous loss. Importantly, the proposed multivariate Pareto enjoys essential level of analytic tractability. This has ensured that 1.) numerous links to certain nowadays existing probabilistic models, as well as seemingly useful characteristic results have been proved, 2.) expressions for, e.g., decumulative distribution functions, densities, (joint) moments and regressions have been derived, and 3.) insurance pricing with general economic weighted pricing functionals have been developed. We believe that the multivariate dependent Pareto distribution introduced in this paper is capable of adequate modeling dependent heavy tailed risks with a non-zero probability of simultaneous loss.

References

Arnold, B. C., 1983. *Pareto distributions.* International Cooperative Publ. House, Fairland, MD.

Balkema, A., de Haan, L., 1974. Residual life time at great age. *Annals of Probability* **2**, 792 – 804.

Bauerle, N., Grubel, R., 2005. Multivariate counting processes: copulas and beyond. *ASTIN Bulletin* **35**(2), 379 – 408.

Boucher, J.P., Denuit, M., Guillén, M., 2008. Models of insurance claim counts with time dependence based on generalization of Poisson and negative binomial distributions. *Variance* **2**(1), 135 – 162.

Cai, J., Tan, K.S., 2007. Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measure. *ASTIN Bulletin* **37**(1), 93 – 112.

Centeno, M. L., 2005. Dependent risks and excess of loss reinsurance. *Insurance: Mathematics and Economics* **37**(2), 229 – 238.

CHERUBINI, U., LUCIANO, E., VECCHIATO, W., 2004. *Copula methods in finance*, Wiley, New York.

CHIRAGIEV, A., LANDSMAN, Z., 2007. Multivariate Pareto portfolios: TCE-based capital allocation and divided differences. *Scandinavian Actuarial Journal* **2007**(4), 261 – 280.

CHIRAGIEV, A., LANDSMAN, Z., 2009. Multivariate flexible Pareto model: Dependency structure, properties and characterizations. *Statistics and Probability Letters* **79**, 1733 – 1743.

FREES, E.W., VALDEZ, E.A., 1998. Understanding relationships using copulas. *North American Actuarial Journal* **2**, 1 – 25.

FURMAN, E., 2008. On a multivariate gamma distribution. *Statistics and Probability Letters* **78**(15), 2353 – 2360.

FURMAN, E., LANDSMAN, Z., 2008. Economic capital allocations for non-negative portfolios of dependent risks. *ASTIN Bulletin* **38**(2), 601 – 619.

FURMAN, E., LANDSMAN, Z., 2009. Multivariate Tweedie distributions and some related capital-at-risk analysis. *Insurance: Mathematics and Economics*, revised and resubmitted.

FURMAN, E., ZITIKIS, R., 2008A. Weighted premium calculation principles. *Insurance: Mathematics and Economics* **42**(1), 459 – 465.

FURMAN, E., ZITIKIS, R., 2008B. Weighted risk capital allocations. *Insurance: Mathematics and Economics* **43**(2), 263 – 270.

FURMAN, E., ZITIKIS, R., 2009. *General Stein-type Covariance Decompositions with Applications to Insurance and Finance.* Actuarial Research Group (ARG) Technical Report No 2009 – 3, University of Western Ontario, London, Ontario, Canada.

HANAGAL, D. D., 1996. A multivariate Pareto distribution. *Communications in Statistics - Theory and Methods* **25**, 1471 – 1488.

KOTZ, S., BALAKRISHNAN, N., JOHNSON, N. L., 2000. *Continuous Multivariate Distributions.* Vol. 1. Models and applications. Wiley, New York.

MARSHALL, A. W., OLKIN, I., 1967. A multivariate exponential distribution. *Journal of American Statistical Association* **62**(317), 30 – 44.

MULIERE, P., SCARSINI, M., 1987. Characterization of a Marshall-Olkin type class of distributions. *Annals of the Institute of Statistical Mathematics* **39**, 429 – 441.

PARETO, V., 1897. *Cours d'Economie Politique.* Paris: Rouge et Cie.

Pfeifer, D., Nešlehová, J., 2004. Modeling and generating dependent risk processes for IRM and DFA, *ASTIN Bulletin* **34**(2), 333 – 360.

Pickands, J., 1975. Statistical inference using extreme order statistics. *The Annals of Statistics* **3**, 119 – 131.

Roger, J., Laeven, A., Goovaerts, M. J., Hoedemakers, T., 2005. Some asymptotic results for sums of dependent random variables with actuarial applications. *Insurance: Mathematics and economics* **37**(2), 154 – 172.

Song, P., 2000. Multivariate dispersion models generated from Gaussian copula. *Scandinavian Journal of Statistics* **27**, 305 – 320.

Tsanakas, A., 2008. Risk measurement in the presence of background risk. *Insurance: Mathematics and Economics* **42**(2): 520 – 528.

Vernic, R., 1997. On the bivariate generalized Poisson distribution. *ASTIN Bulletin* **27**(1), 27 – 31.

Vernic, R., 2000. A multivariate generalization of the generalized Poisson distribution. *ASTIN Bulletin* **30**(1), 57 – 67.

Vernic, R., 2009. Tail conditional expectation for the multivariate Pareto distribution of the second kind: Another approach. *Methodology and Computing in Applied Probability*, to appear.

Appendix A. Proofs

Proof of Proposition 2.1. By definition we have that

$$\begin{aligned}\bar{F}_{\mathbf{X}}(\mathbf{x}) &= \mathbf{P}\left[\cap_{j=1}^{n} X_j > x_j\right] = \mathbf{P}\left[\cap_{j=1}^{n} \min\left(\sigma_j Y_0 + \mu_j, Y_j\right) > x_j\right] \\ &= \mathbf{P}\left[\cap_{j=1}^{n} \left(Y_j > x_j \cap Y_0 > (x_j - \mu_j)/\sigma_j\right)\right] \\ &= \mathbf{P}\left[\left(\cap_{j=1}^{n} Y_j > x_j\right) \cap Y_0 > \max_{j=1,\ldots,n} (x_j - \mu_j)/\sigma_j\right].\end{aligned}$$

Hence, by independence of $Y_0,\ Y_1, \ldots, Y_n$, we obtain that

$$\bar{F}_{\mathbf{X}}(\mathbf{x}) = \bar{F}_{Y_0}\left(\max_{j=1,\ldots,n} \frac{x_j - \mu_j}{\sigma_j}\right) \prod_{j=1}^{n} \bar{F}_{Y_j}(x_j),$$

which completes the proof. $\square$

Proof of Proposition 2.4. We note that, for an $(n+1)$ variate vector $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_n)'$, the unconditional ddf of $\mathbf{Z}$ is

$$\overline{F}_{\mathbf{Z}}(z_1, \ldots, z_n) = \mathbf{E}\left[\overline{F}_{\mathbf{Z}|(\boldsymbol{\Lambda},\Lambda_0)}(z_1, \ldots, z_n|(\boldsymbol{\Lambda}, \Lambda_0))\right]$$
$$= \int_{\mathbf{R}_+^{n+1}} \overline{F}_{\mathbf{Z}|(\boldsymbol{\Lambda},\Lambda_0)}(z_1, \ldots, z_n|\boldsymbol{\lambda}) \prod_{i=0}^{n} f_{\Lambda_i}(\lambda_i) d\boldsymbol{\lambda}, \tag{A.1}$$

where

$$\overline{F}_{\mathbf{Z}|(\boldsymbol{\Lambda},\Lambda_0)}(z_1, \ldots, z_n|\boldsymbol{\lambda}) = \exp\left\{-\lambda_1 z_1 - \cdots - \lambda_n z_n - \lambda_0 \max(z_1, \ldots, z_n)\right\}, \tag{A.2}$$

and

$$f_{\Lambda_i}(\lambda_i) = e^{-\lambda_i} \frac{\lambda_i^{\alpha_i - 1}}{\Gamma(\alpha_i)}, \text{ for } i = 0, 1, \ldots, n. \tag{A.3}$$

Substituting (A.2) and (A.3) into (A.1) and compleeting the latter to $(n+1)$ gamma densities completes the proof. □

Proof of Proposition 2.6. Let $\mathbf{Z} \backsim Pa_n(II)(\mathbf{0}, \mathbf{1}, \boldsymbol{\alpha}, \alpha_0)$, for which the ddf is given by

$$\begin{aligned} \bar{F}_{\mathbf{X}}(z_1, \ldots, z_n) &= \mathbf{P}[Z_1 > z_1, \ldots, Z_n > z_n] \\ &= \left(1 + \max_{j=1,\ldots,n} z_j\right)^{-\alpha_0} \prod_{j=1}^{n} (1 + z_j)^{-\alpha_j}, \quad z_j > 0, j = 1, ..., n. \end{aligned}$$

Further, let G_{n+1} denote an $(n+1)$ variate gamma cdf with independent but not identically distributed univariate margins $Ga(\alpha_i, 1)$, $i = 0, 1, \ldots, n$, and let $\tilde{\mathbf{z}} = (z_1, \ldots, z_n, \max_{j=1,\ldots,n} z_j)'$. Also, denote by $\mathcal{F}(\tilde{\mathbf{z}})$ the $(n+1)$ variate Laplace transform of G_{n+1} evaluated at $\tilde{\mathbf{z}}$. Then we easily observe that, for $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_n)'$,

$$\mathcal{F}(\tilde{\mathbf{z}}) = \bar{F}_{\mathbf{Z}}(z_1, \ldots, z_n) = \int_{\mathbf{R}_+^{n+1}} \exp\left\{-\lambda_1 z_1 - \cdots - \lambda_n z_n - \lambda_0 \max_{j=1,\ldots,n} z_j\right\} dG_{n+1}(\boldsymbol{\lambda}), \tag{A.4}$$

where G_{n+1} is unique because of the corresponding uniqueness of the Laplace transform.

On the other hand, equation (A.4) establishes the desired mixture representation of $\mathbf{Z}$ since for general $\boldsymbol{\Xi}$ and $\boldsymbol{\Lambda}$, it holds that

$$\bar{F}_{\mathbf{Z}}(z_1, \ldots, z_n) = \int_{\mathbf{R}_+^{n+1}} \overline{F}_{\boldsymbol{\Xi}|\boldsymbol{\Lambda}}(z_1, \ldots, z_n|\boldsymbol{\lambda}) dF_{\boldsymbol{\Lambda}}(\boldsymbol{\lambda}),$$

which thus completes the proof. □

Proof of Proposition 2.7. For $j = 1, \ldots, n$, we have that

$$\begin{aligned}\bar{F}_{X_j}(x_j) &= \mathbf{P}[X_j > x_j] = \mathbf{P}[\min(\sigma_j Y_0 + \mu_j, Y_j) > x_j] \\ &= \mathbf{P}[Y_j > x_j \cap Y_0 > (x_j - \mu_j)/\sigma_j] \\ &= \left(1 + \frac{x_j - \mu_j}{\sigma_j}\right)^{-(\alpha_0 + \alpha_j)}, \ x_j > \mu_j,\end{aligned}$$

which completes the proof. □

Proof of Proposition 2.8. Let $\mathbf{Y} = (Y_0, Y_1, \ldots, Y_n)'$ be an $(n+1)$ variate random vector with independent univariate Pareto margins as in Definition 2.1. Also, let $\mathcal{A}_k$ denote the desired event, and let $U_0 = Y_0,\ U_i = (Y_i - \mu_i)/\sigma_i,\ i = 1, \ldots, n.$ Then the U_i's are independent and $Pa(II)(0, 1, \alpha_i)$ distributed, and we have that

$$\mathcal{A}_k = \{\min(U_0, U_{j_1}) = \cdots = \min(U_0, U_{j_k})\},$$

which is equivalent to saying that

$$\mathcal{A}_k = \{U_{j_1} \geq U_0, \ldots, U_{j_k} \geq U_0\} \cup \{U_0 > U_{j_1} = \cdots = U_{j_k}\}.$$

Thus, since

$$\mathbf{P}[U_0 > U_{j_1} = \cdots = U_{j_k}] = 0,$$

we obtain that

$$\mathbf{P}[\mathcal{A}_k] = \mathbf{P}[U_{j_1} \geq U_0, \ldots, U_{j_k} \geq U_0].$$

Further, for $D = \{(u_0, \ldots, u_k) : u_1 \geq u_0, \ldots, u_k \geq u_0, u_0 > x\}$, $x > 0$ and $\mathbf{u} = (u_0, u_1, \ldots, u_k)$, we have that

$$\begin{aligned}&\mathbf{P}[\mathcal{A}_k \cap \{U_0 > x\}] \\ &= \mathbf{P}[U_{j_1} \geq U_0, \ldots, U_{j_k} \geq U_0, U_0 > x] = \int_D \prod_{i=1}^{k} f_{U_{j_i}}(u_i)\, f_{U_0}(u_0)\, d\mathbf{u} \\ &= \int_x^\infty f_{U_0}(u_0) \int_{u_0}^\infty f_{U_{j_1}}(u_1) \ \ldots \int_{u_0}^\infty f_{U_{j_k}}(u_k)\, d\mathbf{u} = \int_x^\infty f_{U_0}(u_0) \prod_{i=1}^{k} \bar{F}_{U_{j_i}}(u_0)\, du_0 \\ &= \alpha_0 \int_x^\infty (1 + u_0)^{-\left(\alpha_0 + \sum_{i=1}^k \alpha_{j_i} + 1\right)} du_0 \\ &= \frac{\alpha_0}{\alpha_0 + \sum_{i=1}^k \alpha_{j_i}} (1 + x)^{-\left(\alpha_0 + \sum_{i=1}^k \alpha_{j_i}\right)}, \ x \geq 0. \qquad \text{(A.5)}\end{aligned}$$

Substituting $x = 0$ in (A.5) completes the proof. □

Proof of Proposition 3.1. It is clear that

$$\bar{F}_{\mathbf{X}}(\mathbf{x}) = \mathbf{P}\left[X_1 > x_1, X_2 > x_2 |\, \mathcal{A}\right] \mathbf{P}\left[\mathcal{A}\right] + \mathbf{P}\left[X_1 > x_1, X_2 > x_2 |\, \bar{\mathcal{A}}\right] \mathbf{P}\left[\bar{\mathcal{A}}\right],$$

where $\mathcal{A} = \{(X_1 - \mu_1)/\sigma_1 = (X_2 - \mu_2)/\sigma_2\}$. Similarly to the proof of Proposition 2.8, let $U_j = (Y_j - \mu_j)/\sigma_j,\ j = 1, 2,\ U_0 = Y_0$. The singular part of $\bar{F}_{\mathbf{X}}$ is then given by

$$\begin{aligned}
(1-a)\,\bar{F}_s(x_1, x_2) &= \mathbf{P}[X_1 > x_1 \cap X_2 > x_2 \cap \mathcal{A}] \\
&= \mathbf{P}\left[\frac{X_1 - \mu_1}{\sigma_1} = \frac{X_2 - \mu_2}{\sigma_2} > \max_{j=1,2} \frac{x_j - \mu_j}{\sigma_j}\right] \\
&= \mathbf{P}\left[\min\left(Y_0, \frac{Y_1 - \mu_1}{\sigma_1}\right) = \min\left(Y_0, \frac{Y_2 - \mu_2}{\sigma_2}\right) > \max_{j=1,2} \frac{x_j - \mu_j}{\sigma_j}\right] \\
&= \mathbf{P}\left[U_1 \geq U_0,\ U_2 \geq U_0,\ U_0 > \max_{j=1,2} \frac{x_j - \mu_j}{\sigma_j}\right],
\end{aligned}$$

which after using (A.5), becomes

$$(1-a)\,\bar{F}_s(x_1, x_2) = \frac{\alpha_0}{\alpha_+}\left(1 + \max_{j=1,2} \frac{x_j - \mu_j}{\sigma_j}\right)^{-\alpha_+}, \quad \frac{x_j - \mu_j}{\sigma_j} \geq 0,\ j = 1,\ 2,$$

and the expression for $\bar{F}_s$ readily follows with $a = 1 - \alpha_0/\alpha_+$. Certainly, the expression for $\bar{F}_{ac}$ results immediately from equation (3.2) as $\bar{F}_{ac} = \frac{\alpha_+}{\alpha_1+\alpha_2}\left(\bar{F}_{\mathbf{X}} - \frac{\alpha_0}{\alpha_+}\bar{F}_s\right)$, which thus completes the proof. □

Proof of Proposition 3.2. We first note in passing that the conditions $\alpha_{0j} > 1$ and $\alpha_{0j} > 2$ assure the finiteness of $\mathbf{E}[X_j]$ and $\mathbf{Var}[X_j], j = 1, 2$, respectively. We further derive the covariance for the standardized random pair $\mathbf{Z} = (Z_1, Z_2) \backsim Pa_2(II)(\mathbf{0}, \mathbf{1}, \boldsymbol{\alpha}, \alpha_0)$, from which $\mathbf{Cov}[X_1, X_2]$ readily follows due to Note 2.2.

Given a function $g : \mathbf{R}_+^2 \to [0, \infty)$, such that $\mathbf{E}[g(X_1, X_2)] < \infty$, and recalling the decomposition

$$\bar{F}_{\mathbf{Z}}(x_1, x_2) = a\bar{F}_{ac,\mathbf{Z}}(x_1, x_2) + (1-a)\bar{F}_{s,\mathbf{Z}}(x_1, x_2),$$

we have, for $\mathbf{x} = (x_1, x_2)'$, that

$$\begin{aligned}
\mathbf{E}\left[g(Z_1, Z_2)\right] &= \int_0^\infty\!\!\int_0^\infty g(\mathbf{x})\, dF_{\mathbf{Z}}(\mathbf{x}) \\
&= \int_0^\infty\!\!\int_0^\infty g(\mathbf{x})\, da F_{ac,\mathbf{Z}}(\mathbf{x}) + \int_0^\infty\!\!\int_0^\infty g(\mathbf{x})\, d(1-a) F_{s,\mathbf{Z}}(\mathbf{x}) \\
&= \iint\limits_{x_1<x_2} g(\mathbf{x})\, a f_{ac,\mathbf{Z}}(\mathbf{x})\, d\mathbf{x} + \iint\limits_{x_1>x_2} g(\mathbf{x})\, a f_{ac,\mathbf{Z}}(\mathbf{x})\, d\mathbf{x} + \iint\limits_{x_1=x_2} g(\mathbf{x})\, d\frac{\alpha_0}{\alpha_+} F_{s,\mathbf{Z}}(\mathbf{x}).
\end{aligned}$$

In particular, for the function $g(x_1,x_2)=(x_1-\mathbf{E}[Z_1])(x_2-\mathbf{E}[Z_2])$, we have that

$$\begin{aligned}
&\mathbf{Cov}\left[Z_1,Z_2\right]=\mathbf{E}\left[\left(Z_1-\mathbf{E}[Z_1]\right)\left(Z_2-\mathbf{E}[Z_2]\right)\right]\\
=&\int\limits_0^\infty (x_2-\mathbf{E}[Z_2])\left(\int\limits_0^{x_2}(x_1-\mathbf{E}[Z_1])\,af_{ac,\mathbf{Z}}\left(\mathbf{x}\right)dx_1+\int\limits_{x_2}^\infty (x_1-\mathbf{E}[Z_1])\,af_{ac,\mathbf{Z}}\left(\mathbf{x}\right)dx_1\right)dx_2\\
+&\int_0^\infty (x_2-\mathbf{E}[Z_2])\,(x_2-\mathbf{E}[Z_1])\,d\frac{\alpha_0}{\alpha_+}\left(1-(1+x_2)^{-\alpha_+}\right).
\end{aligned}$$

Further, substituting the expression for $af_{ac,\mathbf{Z}}$ from Note 3.2, recalling that $\mathbf{E}[Z_j]=(\alpha_{0j}-1)^{-1}$, and setting $y_j=1+x_j$, where $j=1,2$, the covariance becomes

$$\begin{aligned}
\mathbf{Cov}\left[Z_1,Z_2\right] = &\int\limits_1^\infty\left(y_2-\frac{\alpha_{02}}{\alpha_{02}-1}\right)\left(\alpha_1\alpha_{02}y_2^{-(\alpha_{02}+1)}\int\limits_1^{y_2}\left(y_1-\frac{\alpha_{01}}{\alpha_{01}-1}\right)y_1^{-(\alpha_1+1)}dy_1\right.\\
&+\left.\alpha_2\alpha_{01}y_2^{-(\alpha_2+1)}\int\limits_{y_2}^\infty\left(y_1-\frac{\alpha_{01}}{\alpha_{01}-1}\right)y_1^{-(\alpha_{01}+1)}dy_1\right)dy_2\\
&+\alpha_0\int\limits_1^\infty\left(y_2-\frac{\alpha_{02}}{\alpha_{02}-1}\right)\left(y_2-\frac{\alpha_{01}}{\alpha_{01}-1}\right)y_2^{-(\alpha_++1)}dy_2.
\end{aligned}$$

After evaluating the inner integrals, the covariance further reduces to

$$\begin{aligned}
&\mathbf{Cov}\left[Z_1,Z_2\right]\\
=&\frac{\alpha_0}{(\alpha_1-1)(\alpha_{01}-1)}\int\limits_1^\infty\left(\alpha_{02}\left(y_2^{-\alpha_{02}}-\frac{\alpha_{02}y_2^{-\alpha_{02}-1}}{\alpha_{02}-1}\right)+(1-\alpha_+)\left(y_2^{-\alpha_++1}-\frac{\alpha_{02}y_2^{-\alpha_+}}{\alpha_{02}-1}\right)\right)dy_2,
\end{aligned}$$

which, for $\alpha_+>2$, yields

$$\mathbf{Cov}\left[Z_1,Z_2\right]=\frac{\alpha_0}{(\alpha_{01}-1)(\alpha_{02}-1)(\alpha_+-2)},$$

from where the stated formula follows, since due to Note 2.2, $\mathbf{Cov}\left[X_1,X_2\right]=\sigma_1\sigma_2\mathbf{Cov}\left[Z_1,Z_2\right]$.

Certainly, the expression for $\mathbf{Corr}\left[X_1,X_2\right]$ results immediately, and thus completes the proof. □

Proof of Proposition 3.3. The absolutely continuous component of the distribution of X_1 given $X_2=x_2$ has density

$$a'f_{ac}(x_1|x_2)=\frac{af_{ac}\left(\mathbf{x}\right)}{f_{X_2}\left(x_2\right)}$$

$$=\begin{cases}\dfrac{\alpha_{01}\alpha_2}{\alpha_{02}\sigma_1}\left(1+\dfrac{x_1-\mu_1}{\sigma_1}\right)^{-(\alpha_{01}+1)}\left(1+\dfrac{x_2-\mu_2}{\sigma_2}\right)^{\alpha_0}, & \dfrac{x_1-\mu_1}{\sigma_1}>\dfrac{x_2-\mu_2}{\sigma_2}>0\\[2ex] \dfrac{\alpha_1}{\sigma_1}\left(1+\dfrac{x_1-\mu_1}{\sigma_1}\right)^{-(\alpha_1+1)}, & \dfrac{x_2-\mu_2}{\sigma_2}>\dfrac{x_1-\mu_1}{\sigma_1}>0\end{cases},$$

where the density of X_2, denoted f_{X_2}, results from Proposition 2.7.

Then the discrete component corresponding to $\{(x_1-\mu_1)/\sigma_1 = (x_2-\mu_2)/\sigma_2\}$ is $1-a'$ with a' given by

$$a' = \int_{\mu_1}^{\infty} \frac{a f_{ac}(\mathbf{x})}{f_{X_2}(x_2)} dx_1 = 1 - \frac{\alpha_0}{\alpha_{02}} \left(1 + \frac{x_2-\mu_2}{\sigma_2}\right)^{-\alpha_1}.$$

Consequently, the ddf of the conditional distribution results immediately by integration. However, special attention must be paid to the case $\frac{x_2-\mu_2}{\sigma_2} \geq \frac{x_1-\mu_1}{\sigma_1}$, when we have that

$$\bar{F}_{X_1|X_2}(x_1|x_2) = \int_{x_1}^{\mu_1+\frac{\sigma_1}{\sigma_2}(x_2-\mu_2)} a' f_{ac}(x|x_2)dx + \int_{\mu_1+\frac{\sigma_1}{\sigma_2}(x_2-\mu_2)}^{\infty} a' f_{ac}(x|x_2)dx + (1-a'),$$

which completes the proof. $\square$

Proof of Corollary 3.1. We give a proof for $\mathbf{Z} = (Z_1, Z_2) \backsim Pa_2(II)(\mathbf{0}, \mathbf{1}, \boldsymbol{\alpha}, \alpha_0)$, and $\alpha_1 \neq 1$. The stated formula then follows as a result of Note 2.2. Namely, for $x_2 > 0$, we have that

$$\begin{aligned} \mathbf{E}[Z_1|Z_2 = x_2] &= -\int_0^{\infty} x_1 d\bar{F}_{Z_1|Z_2}(x_1|x_2) = \alpha_1 \int_0^{x_2} x_1 (1+x_1)^{-\alpha_1-1} dx_1 \\ &+ \frac{\alpha_{01}\alpha_2}{\alpha_{02}} (1+x_2)^{\alpha_0} \int_{x_2}^{\infty} x_1 (1+x_1)^{-\alpha_{01}-1} dx_1 + \frac{\alpha_0}{\alpha_{02}} x_2 (1+x_2)^{-\alpha_1} \\ &= \frac{1}{\alpha_1 - 1} \left(1 + \frac{\alpha_0(1-\alpha_+)}{\alpha_{02}(\alpha_{01}-1)} (1+x_2)^{-(\alpha_1-1)}\right). \end{aligned}$$

The remaining case for $\alpha_1 = 1$ is then accomplished in the same fashion. This completes the proof. $\square$