ON THE EVALUATION OF MULTIVARIATE COMPOUND DISTRIBUTIONS WITH CONTINUOUS SEVERITY DISTRIBUTIONS AND SARMANOV'S COUNTING DISTRIBUTION

BY

MAISSA TAMRAZ AND RALUCA VERNIC

ABSTRACT

In this paper, we present closed-type formulas for some multivariate compound distributions with multivariate Sarmanov counting distribution and independent Erlang distributed claim sizes. Further on, we also consider a type-II Pareto dependency between the claim sizes of a certain type. The resulting densities rely on the special hypergeometric function, which has the advantage of being implemented in the usual software. We numerically illustrate the applicability and efficiency of such formulas by evaluating a bivariate cumulative distribution function, which is also compared with the similar function obtained by the classical recursion-discretization approach.

KEYWORDS

Multivariate compound model, sarmanov's multivariate discrete distribution, Erlang distribution, type-II Pareto multivariate distribution, hypergeometric function.

1. INTRODUCTION

Used to model the aggregate claims of a portfolio, the univariate collective model is represented as

$$S = \sum_{l=0}^{N} X_l, \tag{1}$$

where N is the random variable (r.v.) number of claims and $(X_l)_{l\geq 1}$ are the corresponding nonnegative r.v.s claim sizes with $X_0 = 0$. The classical hypotheses that provide the tractability of this model are independent, identically distributed (i.i.d.) discrete claim sizes, also independent of N. There is a large amount of literature related to the evaluation of the compound distribution of S under these assumptions, see, e.g., Klugman *et al.* (1998) for different methods,

Astin Bulletin 48(2), 841-870. doi: 10.1017/asb.2017.46

and Sundt and Vernic (2009) for a survey of the recursive methods. However, in practice, the claim sizes (severity) distribution is rather of continuous type, hence the usual approach consists in two steps: in the first step, the severity distribution is discretized, while in the second step, a specific method is applied to the resulting discrete compound distribution. In this case, a special attention must be paid to the choice of the discretization span: a large span can generate important errors, while a very small span can lead to a very long running time, especially in the multivariate case. In this respect, it is unfortunate that closed-type formulas for compound distributions with continuous type claim sizes are so scarce; in the univariate case, apart the Gamma severity distribution (which also includes the well-known exponential case) leading to the so-called Tweedie compound distribution (see, e.g., Dunn and Smyth, 2005), we mention the recent work of Sarabia *et al.* (2016), who went even further on by considering a Pareto-type dependency between the aggregated claim sizes.

In this paper, we propose closed-type formulas for some multivariate compound distributions with Sarmanov counting distribution and Erlang severity distributions; furthermore, inspired by Sarabia *et al.* (2016), we also include some dependency between the claim sizes. Our formulas are expressed mainly in terms of the special hypergeometric function already implemented in the existing mathematical software, hence making the related calculations numerically feasible without involving other techniques.

More precisely, we deal with a multivariate extension of model (1), i.e., for $m \geq 2$, we consider

$$(S_1, \ldots, S_m) = \left(\sum_{l=0}^{N_1} X_{1l}, \ldots, \sum_{l=0}^{N_m} X_{ml} \right), \tag{2}$$

where N_j denotes the number of claims of type j and $\left(X_{jl}\right)_{l \geq 1}$ the corresponding claim sizes, where by convention, $X_{j0} = 0$, $0 \leq j \leq m$. This model corresponds to the situation where we have m different types of claims generated by some related events, hence the claim numbers $\left(N_j\right)_{j=1}^{m}$ are dependent. The model has been studied mostly under the assumptions that the claims of type j are i.i.d., independent of the claim numbers and independent of the claims of type k, $\forall j \neq k$ (see, e.g., Sundt and Vernic, 2009, Jin and Ren, 2014 or Robe-Voinea and Vernic, 2016). We shall call by "inside-type independency" the independency assumption between the claims of same type, while by "between-types independency" we designate the independency assumptions between claims of different types. Then, similarly with Sarabia *et al.* (2016), we shall relax the inside-type independency condition by considering a certain type of dependency in each set of claims $\left(X_{jl}\right)_{l \geq 1}$. Note that in the univariate case, dependency between the individual risks has already been considered especially in the individual model, see, e.g., Goovaerts and Dhaene (1996), Genest *et al.* (2003), Denuit *et al.* (2005) and the references therein.

Regarding the claim sizes distributions, we choose the Erlang distribution in the inside-type independency case and, when relaxing this condition, the same multivariate type-II Pareto distribution as in Sarabia *et al.* (2016). This choice has been driven by the interest in obtaining closed formulas numerically computable with the existing software, but also by the fact that both distributions have been intensively studied in the actuarial literature lately; for the Erlang distribution, see, e.g., Willmot and Lin (2011) or Willmot and Woo (2015), while for the multivariate type-II Pareto distribution, see Asimit *et al.* (2013) or the generalizations in Asimit *et al.* (2010) and Guillén *et al.* (2013).

Therefore, this paper is structured as follows: in Section 2, we introduce some notation and recall several special functions and distributions that will be used in the sequel. In Section 3, we present closed-type formulas for univariate compound distributions with Erlang severity distribution, while in Section 4 we extend these formulas to multivariate compound distributions with Sarmanov counting distribution; moreover, Section 4 is divided into two subsections corresponding to the cases with and without inside-type independency. In the bivariate case, a special attention is paid to the correlation coefficient of the resulting compound distribution, which is expressed in terms of the correlation coefficient of the original counting distribution, and results smaller than the last one.

To illustrate the applicability, efficiency and importance of the derived formulas, in Section 5, we present a numerical example in which we compare the cumulative distribution functions (cdfs) obtained for a particular bivariate compound distribution by using the closed-type formulas and the usual recursion-discretization approach. The paper ends with some conclusions and future work, followed by an Appendix containing the proofs.

2. PRELIMINARIES

2.1. Notation, definitions and useful formulas

In connection with the univariate collective model (1), we denote the probability mass function (pmf) of the discrete r.v. N by p, while h denotes the probability density function (pdf) of the claim sizes, which are assumed to be positive, continuous and identically distributed (i.d.), not necessarily independent. Therefore, the distribution of S is called *compound* with counting distribution p and severity distribution h; we denote it by $p \vee h$. Then, letting $h^{(n)}$ denote the pdf of the sum $\sum_{i=0}^{n} X_i$, where $h^{(0)}(x) = \begin{cases} 1, x = 0 \\ 0, \text{otherwise} \end{cases}$, it holds that (see Sarabia *et al.*, 2016)

$$(p \vee h)(x) = \sum_{n=0}^{\infty} p(n)\, h^{(n)}(x), \; x \geq 0, \tag{3}$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

hence $(p \vee h)(0) = p(0)$. Under the classical assumption that the claim amounts are also independent, this formula reduces to the well-known one

$$(p \vee h)(x) = \sum_{n=0}^{\infty} p(n) h^{*n}(x), \tag{4}$$

where h^{*n} denotes the n-fold convolution of h.

To simplify the writing in the multivariate case, we denote $\mathbf{X} = (X_1, \ldots, X_m)$ or $\mathbf{x} = (x_1, \ldots, x_m)$, $\overline{1, m} = \{1, 2, \ldots, m\}$; moreover, $\mathbf{0}$ denotes the 0-vector, while $\mathbf{x} - \mathbf{y}$ and $\mathbf{x} \geq \mathbf{y}$ are considered componentwise. In what concerns the model (2), the pmf of the random vector consisting of the (dependent) claim numbers $(N_1, \ldots, N_m)$ is still denoted by p, while h_j denotes the pmf of the i.d. continuous positive claim amounts of type j, and $\mathbf{h} = (h_1, \ldots, h_m)$. With this notation, assuming that both inside-type and between-types independency conditions hold, the compound pdf $(p \vee \mathbf{h})$ of $\mathbf{S} = (S_1, \ldots, S_m)$ can be written as

$$(p \vee \mathbf{h})(\mathbf{x}) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} p(\mathbf{n}) \prod_{j=1}^{m} h_j^{*n_j}(x_j), \ \mathbf{x} \geq \mathbf{0}. \tag{5}$$

However, relaxing the inside-type independency assumption while keeping the between-types independency condition, the distribution of $\mathbf{S}$ becomes (we omit the proof being similar with the one in Sarabia *et al.*, 2016)

$$(p \vee \mathbf{h})(\mathbf{x}) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} p(\mathbf{n}) \prod_{j=1}^{m} h_j^{(n_j)}(x_j), \ \mathbf{x} \geq \mathbf{0}. \tag{6}$$

It follows that $(p \vee \mathbf{h})(\mathbf{0}) = p(\mathbf{0})$, and, marginally, e.g.,

$$(p \vee \mathbf{h})(x_1, 0, \ldots, 0) = \sum_{n_1=1}^{\infty} p(n_1, 0, \ldots, 0) h_1^{(n_1)}(x_1), x_1 > 0,$$

$$(p \vee \mathbf{h})(0, x_2, \ldots, x_m) = \sum_{n_2=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} p(0, n_2, \ldots, n_m) \times \prod_{j=2}^{m} h_j^{(n_j)}(x_j), x_2, \ldots, x_m > 0.$$

Note that in both univariate and multivariate cases, the formulas (3) and (6) can be used as definitions for a more general compounding operation involving a discrete function $p : \mathbb{N}^m \to \mathbb{R}, m \geq 1$, which is not necessarily a pmf (see also, Vernic, 2017); however, we keep the assumption that the functions h and h_js are pdfs.

We shall now recall several special functions. The Laplace transform of a r.v. X is defined by $\mathcal{L}_X(t) = \mathbb{E}\left[e^{-tX}\right]$.

Spence's function or dilogarithm, which is a particular case of the polylogarithm function, is defined for $|t| < 1$ by the power series $Li_2(t) = \sum_{k=1}^{\infty} t^k/k^2$.

We let

$$I_0(x) = \sum_{n=0}^{\infty} (x/2)^{2n} \frac{1}{(n!)^2},\ I_1(x) = \sum_{n=0}^{\infty} (x/2)^{2n+1} \frac{1}{n!\,(n+1)!}$$

denote the modified Bessel functions of first and, respectively, second kind.

Based on the Pochhammer symbol $(a)_{(n)} = a(a+1)\times\cdots\times(a+n-1)$, $n \geq 1$, $(a)_{(0)} = 1$, the generalized hypergeometric function ${}_rF_q$ is defined by

$${}_rF_q\left(\{a_1,\ldots,a_r\},\{b_1,\ldots,b_q\};z\right) = 1 + \sum_{n=1}^{\infty} \frac{(a_1)_{(n)}\times\cdots\times(a_r)_{(n)}}{(b_1)_{(n)}\times\cdots\times\left(b_q\right)_{(n)}}\frac{z^n}{n!}. \quad (7)$$

We shall need the following result (its proof is given in the Appendix). By convention, an empty product equals 1.

Lemma 2.1. *For $n \in \mathbb{N}$ and $k \in \mathbb{N}^*$, it holds that*

$$i)\ \frac{1}{((n+1)k-1)!\,(n+1)} = \frac{1}{(k-1)!k^{nk}\left(\frac{k+1}{k}\right)_{(n)}\left(\frac{k+2}{k}\right)_{(n)}\times\cdots\times\left(\frac{k+(k-1)}{k}\right)_{(n)}(2)_{(n)}}, \quad (8)$$

$$ii)\ \frac{1}{(nk)!} = \frac{1}{n!k^{nk}\left(\frac{1}{k}\right)_{(n)}\left(\frac{2}{k}\right)_{(n)}\times\cdots\times\left(\frac{k-1}{k}\right)_{(n)}}. \quad (9)$$

Note that due to the convention, when $k = 1$ formula (8) becomes $\frac{1}{n!(n+1)} = \frac{1}{(2)_{(n)}}$, while formula (9) yields the identity $\frac{1}{n!} = \frac{1}{n!}$.

2.2. Some distributions

We shall now recall some distributions needed in the sequel.

2.2.1. *Univariate distributions.* In the discrete case, we shall use the well-known Poisson distribution $\mathrm{Po}(\mu)$, $\mu > 0$, the negative binomial distribution $\mathrm{NB}(r,q)$, $r > 0$, $q \in (0,1)$, with pmf $\frac{\Gamma(r+n)}{n!\Gamma(r)}q^r(1-q)^n$, $n \in \mathbb{N}$, expected value $\frac{r(1-q)}{q}$ and variance $\frac{r(1-q)}{q^2}$, and the logarithmic distribution $\mathrm{Log}(\theta)$, $\theta \in (0,1)$, with pmf $\frac{-1}{\ln(1-\theta)}\frac{\theta^n}{n}$, $n \geq 1$, expected value $\frac{-\theta}{(1-\theta)\ln(1-\theta)}$ and variance $\frac{-\theta(\theta+\ln(1-\theta))}{(1-\theta)^2\ln^2(1-\theta)}$.

In the continuous case, we recall the Gamma distribution $\mathrm{Ga}(\alpha,\beta)$, $\alpha,\beta > 0$, whose pdf is given by $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$, $x > 0$, expected value $\frac{\alpha}{\beta}$ and variance $\frac{\alpha}{\beta^2}$, where $\Gamma(\cdot)$ denotes the Gamma function; it is well known that the

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

n-fold convolution of $\mathrm{Ga}(\alpha, \beta)$ is still Gamma distributed, i.e., $\mathrm{Ga}(n\alpha, \beta)$. In particular, when $\alpha = k \in \mathbb{N}^*$, the Gamma distribution is called Erlang that we shall denote by $\mathrm{Erlang}(k, \beta)$. Also, $\mathrm{Ga}(1, \beta)$ is the exponential distribution denoted by $\mathrm{Exp}(\beta)$.

Another distribution that we shall encounter in the following is the beta distribution of the second kind, also called beta prime, inverted beta or Pearson type VI distribution (for details on this distribution, see, e.g., Kleiber and Kotz, 2003). The pdf of this distribution is given by $f(x) = \frac{x^{\beta-1}}{B(\alpha,\beta)(1+x)^{\alpha+\beta}}, x > 0$, where $\alpha, \beta > 0$ and $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ denotes the beta function. Introducing also a scale parameter $\sigma > 0$, the pdf becomes

$$f(x) = \frac{x^{\beta-1}}{\sigma^{\beta} B(\alpha, \beta)(1 + x/\sigma)^{\alpha+\beta}}, x > 0.$$

We denote this distribution by $\mathrm{B}_{II}(\beta, \alpha, \sigma)$ and note that it can be obtained as the distribution of the ratio of two independent r.v.s, i.e., as $\sigma \frac{Y}{Z}$, where $Y \sim \mathrm{Ga}(\beta, 1)$ and $Z \sim \mathrm{Ga}(\alpha, 1)$.

Moreover, the ratio $\sigma \frac{Y}{Z}$ of two independent r.v.s, where $\sigma > 0, Z \sim \mathrm{Ga}(\alpha, 1)$ and $Y \sim \mathrm{Exp}(1)$, follows a Pareto distribution $\mathrm{Pa}(\alpha, \sigma)$ with pdf $\frac{\alpha}{\sigma}\left(1 + \frac{x}{\sigma}\right)^{-\alpha-1}, x > 0$, expected value $\frac{\sigma}{\alpha-1}, \alpha > 1$, and variance $\frac{\alpha\sigma^2}{(\alpha-1)^2(\alpha-2)}$, $\alpha > 2$.

2.2.2. *Multivariate type-II Pareto distribution.* Starting from m i.i.d. r.v.s $Y_1, \ldots, Y_m$ exponentially $\mathrm{Exp}(1)$ distributed and independent of the r.v. $Z \sim \mathrm{Ga}(\alpha, 1), \alpha > 0$, the random vector defined by $\mathbf{X} = \left(\sigma \frac{Y_1}{Z}, \ldots, \sigma \frac{Y_m}{Z}\right), \sigma > 0$, follows an m-variate Pareto of type-II distribution with pdf

$$f(\mathbf{x}) = \frac{\Gamma(\alpha + m)}{\Gamma(\alpha)\sigma^m}\left(1 + \frac{1}{\sigma}\sum_{i=1}^{m} x_i\right)^{-\alpha-m}, x_1, \ldots, x_m > 0.$$

We denote this distribution by $\mathrm{Pa}II_m(\alpha, \sigma)$ and note that its marginals are all i.d. $\mathrm{Pa}(\alpha, \sigma)$, the covariance between components is $\mathrm{cov}(X_i, X_j) = \frac{\sigma^2}{(\alpha-1)^2(\alpha-2)}, \alpha > 2, i \neq j$, and the correlation $\rho = \frac{1}{\alpha}$. Sarabia *et al.* (2016) also showed that the sum of the components $\sum_{i=1}^{m} X_i$ follows the beta distribution of the second kind, $\mathrm{B}_{II}(m, \alpha, \sigma)$. For more details on the Pareto distribution, see Arnold (2015).

2.2.3. *Sarmanov's multivariate distribution.* We recall that the random vector $\mathbf{N} = (N_1, \ldots, N_m)$ follows an m-variate discrete Sarmanov distribution with

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

joint pmf given for $\mathbf{n} \in \mathbb{N}^m$ by (see, e.g., Kotz *et al.*, 2000)

$$p(\mathbf{n}) = \left(\prod_{l=1}^{m} p_l(n_l)\right)\left(1 + \sum_{k=2}^{m} \sum_{1 \le j_1 < \ldots < j_k \le m} \omega_{j_1 \ldots j_k} \phi_{j_1}(n_{j_1}) \times \cdots \times \phi_{j_k}(n_{j_k})\right), \tag{10}$$

where $(p_l)_{l=1}^m$ are the marginal pmf-s, $(\phi_j)_{j=1}^m$ are bounded non-constant kernel functions, and the constants $\omega_{j_1 \ldots j_k} \in \mathbb{R}$ are such that the following conditions hold:

$$\sum_{n \in \mathbb{N}} \phi_j(n)\, p_j(n) = 0, \ \forall j \in \overline{1, m}, \text{ and} \tag{11}$$

$$1 + \sum_{k=2}^{m} \sum_{1 \le j_1 < \ldots < j_k \le m} \omega_{j_1 \ldots j_k} \phi_{j_1}(n_{j_1}) \times \cdots \times \phi_{j_k}(n_{j_k}) \ge 0, \ \forall \mathbf{n} \in \mathbb{N}^m. \tag{12}$$

The joint distribution of any subset of marginals of $\mathbf{N}$ is of the same type. The particular forms discussed in the literature for the functions ϕ_j can be unified into a general one satisfying condition (11), i.e.,

$$\phi_j(x) = f_j(x) - \mathbb{E}\left[f_j(N_j)\right], \tag{13}$$

with the functions f_j properly chosen such that $\mathbb{E}\left[f_j(N_j)\right] < \infty$. For simplicity, we denote $E_j := \mathbb{E}\left[f_j(N_j)\right]$. In our study, we shall consider the following particular cases:

1. $f(x) = e^{-\delta x} \Rightarrow \phi(x) = e^{-\delta x} - \mathcal{L}_N(\delta)$; a frequent choice is $\delta = 1$.
2. $f = p \Rightarrow \phi(x) = p(x) - \sum_{n \in \mathbb{N}} p^2(n)$.

Remark 2.1. *In the bivariate case where* $p(\mathbf{n}) = \prod_{i=1}^{2} p_i(n_i) \times \left(1 + \omega \prod_{i=1}^{2} \phi_i(n_i)\right), \mathbf{n} \in \mathbb{N}^2$, *condition (12) yields the following range of* ω*:*

$$\max\left\{-\frac{1}{m_1 m_2}, -\frac{1}{M_1 M_2}\right\} \le \omega \le \min\left\{-\frac{1}{m_1 M_2}, -\frac{1}{m_2 M_1}\right\},$$

where $m_i = \min_{n \in \mathbb{N}} \phi_i(n)$, $M_i = \max_{n \in \mathbb{N}} \phi_i(n)$. *From these limits, we can also obtain the correlation range, where the correlation coefficient is given by*

$$\rho = \omega \frac{\mathbb{E}\left[N_1 \phi_1(N_1)\right] \mathbb{E}\left[N_2 \phi_2(N_2)\right]}{\sqrt{Var\left[N_1\right] Var\left[N_2\right]}}. \tag{14}$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

3. Some univariate compound distributions with Erlang severity distribution

Let us consider the univariate compound model (1) with independent claim sizes Erlang (k, β) distributed, $k \in \mathbb{N}^*, \beta > 0$. Then, according to (4), $(p \vee h)(0) = p(0)$ and, for $x > 0$,

$$(p \vee h)(x) = \sum_{n=1}^{\infty} p(n) \frac{\beta^{nk}}{(nk-1)!} x^{nk-1} e^{-\beta x}. \tag{15}$$

Regarding the choice of p, we consider the following three cases for which we express the resulting pdf in terms of the generalized hypergeometric function.

Proposition 3.1. *i. Poisson case: let $N \sim Po(\mu)$. Then $(p \vee h)(0) = e^{-\mu}$ and for $x > 0$,*

$$(p \vee h)(x) = e^{-\mu-\beta x} \frac{\mu (\beta x)^k}{x(k-1)!} \times {}_0F_k\left(\{\}, \left\{\frac{k+1}{k}, \frac{k+2}{k}, \ldots, \frac{2k-1}{k}, 2\right\}; \mu\left(\frac{\beta x}{k}\right)^k\right).$$

ii. Negative binomial case: assuming that $N \sim NB(r, q), r > 0, q \in (0,1)$, we have $(p \vee h)(0) = q^r$, while for $x > 0$,

$$(p \vee h)(x) = \frac{r q^r (1-q)(\beta x)^k e^{-\beta x}}{x(k-1)!} \times {}_1F_k\left(\{r+1\}, \left\{\frac{k+1}{k}, \frac{k+2}{k}, \ldots, \frac{2k-1}{k}, 2\right\}; (1-q)\left(\frac{\beta x}{k}\right)^k\right).$$

iii. Logarithmic case: assume that $N \sim Log(\theta), \theta \in (0,1)$. Then $(p \vee h)(0) = 0$, while for $x > 0$,

$$(p \vee h)(x) = \begin{cases} -\dfrac{e^{-\beta x}}{x \ln(1-\theta)} \left(e^{\theta \beta x} - 1\right), k = 1 \\ -\dfrac{k e^{-\beta x}}{x \ln(1-\theta)} \left[{}_0F_k\left(\{\}, \left\{\dfrac{1}{k}, \dfrac{2}{k}, \ldots, \dfrac{k-1}{k}\right\}; \theta\left(\dfrac{\beta x}{k}\right)^k\right) - 1\right], k \geq 2 \end{cases}.$$

Remark 3.1. *The above proposition can be easily extended to the case where the counting distribution is a mixture with the corresponding components being distributions considered in the proposition. For example, if N follows a mixture of two*

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

Poisson distributions, i.e.,

$$p(n) = qe^{-\mu_1}\frac{\mu_1^n}{n!} + (1-q)\,e^{-\mu_2}\frac{\mu_2^n}{n!}, n \in \mathbb{N}, q \in (0,1), \mu_i > 0, i = 1, 2,$$

then we easily obtain that for $x > 0$,

$$\begin{aligned}(p \vee h)(x) &= \frac{e^{-\beta x}(\beta x)^k}{x(k-1)!} \\ &\times \left[q\mu_1 e^{-\mu_1} {}_0F_k\left(\{\}, \left\{\frac{k+1}{k}, \frac{k+2}{k}, \ldots, \frac{2k-1}{k}, 2\right\}; \mu_1\left(\frac{\beta x}{k}\right)^k\right)\right. \\ &\left. + (1-q)\,\mu_2 e^{-\mu_2} {}_0F_k\left(\{\}, \left\{\frac{k+1}{k}, \frac{k+2}{k}, \ldots, \frac{2k-1}{k}, 2\right\}; \mu_2\left(\frac{\beta x}{k}\right)^k\right)\right].\end{aligned}$$

Remark 3.2. *The distribution obtained in the first case (i) of Proposition 3.1 belongs to the class of Tweedie distributions, which in some cases can be modeled as compound distributions with Poisson counting distribution and Gamma severity distribution, see, e.g., Jorgensen (1997).*

Sarabia *et al.* (2016) presented closed-type formulas for the univariate collective model (1) under the assumption that the claim sizes are dependent according to the multivariate type-II Pareto distribution described in Section 2.2.2. In the following lemma, we recall their results for the same counting distributions considered above.

Lemma 3.1. *If the claim sizes are multivariate type-II Pareto distributed with parameters* (α, σ)*, then*

i. *Poisson case: let* $N \sim Po(\mu)$. *Then,* $(p \vee h)(0) = e^{-\mu}$ *and for* $x > 0$,

$$(p \vee h)(x) = \frac{\alpha\mu e^{-\mu}}{\sigma(1+x/\sigma)^{1+\alpha}}\, {}_1F_1\left(\{1+\alpha\}, \{2\}; \frac{\mu x}{\sigma + x}\right).$$

ii. *Negative binomial case: assuming that* $N \sim NB(r, q), r > 0, q \in (0,1)$, *we have* $(p \vee h)(0) = q^r$, *while for* $x > 0$,

$$(p \vee h)(x) = \frac{r\alpha(1-q)q^r}{\sigma(1+x/\sigma)^{1+\alpha}}\, {}_2F_1\left(\{1+r, 1+\alpha\}, \{2\}; \frac{(1-q)x}{\sigma + x}\right).$$

iii. *Logarithmic case: assume that* $N \sim Log(\theta), \theta \in (0,1)$. *Then,* $(p \vee h)(0) = 0$, *while for* $x > 0$,

$$(p \vee h)(x) = -\frac{1}{x\ln(1-\theta)}\left[\frac{1}{(1+(1-\theta)x/\sigma)^{\alpha}} - \frac{1}{(1+x/\sigma)^{\alpha}}\right].$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

4. Multivariate compound distributions with Sarmanov's counting distribution

In this section, we extend the above results to the multivariate case corresponding to model (2). We assume that the vector number of claims follows Sarmanov's multivariate distribution. Regarding the claim sizes, we first assume the existence of both inside-type and between-types independencies, then we relax the inside-type independency condition similarly to Sarabia *et al.* (2016). In both situations, the following result holds.

Proposition 4.1. *Consider the multivariate compound distribution (6) under the assumption that the multivariate counting distribution p is of Sarmanov type (10). Then the resulting compound distribution also belongs to Sarmanov's class, satisfying for* $\mathbf{s} \geq \mathbf{0}$,

$$(p \vee \mathbf{h})(\mathbf{s}) = \prod_{l=1}^{m} (p_l \vee h_l)(s_l) + \sum_{k=2}^{m} \sum_{1 \leq j_1 < \ldots < j_k \leq m} \omega_{j_1 \ldots j_k} \times \prod_{l=1}^{k} \left(\left(f_{j_l} p_{j_l} \right) \vee h_{j_l} - E_{j_l} \left(p_{j_l} \vee h_{j_l} \right) \right) \left(s_{j_l} \right) \prod_{l=k+1}^{m} \left(p_{j_l} \vee h_{j_l} \right) \left(s_{j_l} \right) \tag{16}$$

$$= \prod_{l=1}^{m} (p_l \vee h_l)(s_l) \left(1 + \sum_{k=2}^{m} \sum_{1 \leq j_1 < \ldots < j_k \leq m} \omega_{j_1 \ldots j_k} \tilde{\phi}_{j_1} \left(s_{j_1} \right) \times \cdots \times \tilde{\phi}_{j_k} \left(s_{j_k} \right) \right), \tag{17}$$

where the indexes $\{j_{k+1}, \ldots, j_m\} = \{1, \ldots, m\} \setminus \{j_1, \ldots, j_k\}$ *and* $\tilde{\phi}_j = \frac{(f_j p_j) \vee h_j}{p_j \vee h_j} - E_j$, $j = \overline{1, m}$.

Remark 4.1. *From the above definition of* $\tilde{\phi}_j$, *it results that* $\tilde{\phi}_j(0) = \frac{(f_j p_j)(0)}{p_j(0)} - E_j = f_j(0) - E_j = \phi_j(0)$, *from where formula (17) yields* $(p \vee \mathbf{h})(\mathbf{0}) = p(\mathbf{0})$ *as expected, and, e.g.,*

$$(p \vee \mathbf{h})(s, 0, \ldots, 0) = (p_1 \vee h_1)(s) \left(\prod_{l=2}^{m} p_l(0) \right) \times \left[1 + \sum_{k=1}^{m-1} \sum_{2 \leq j_1 < \ldots < j_k \leq m} \omega_{1 j_1 \ldots j_k} \tilde{\phi}_1(s) \prod_{l=1}^{k} \phi_{j_l}(0) + \sum_{k=2}^{m-1} \sum_{2 \leq j_1 < \ldots < j_k \leq m} \omega_{j_1 \ldots j_k} \prod_{l=1}^{k} \phi_{j_l}(0) \right], s > 0.$$

As a consequence, the compound distribution $p \vee \mathbf{h}$ *has now continuous and discrete parts (with pmf in* $\mathbf{0}$*).*

In what concerns the correlation, the following proposition deals with the bivariate case.

Proposition 4.2. *The correlation coefficient of the bivariate compound distribution defined by (2) with pdf (17) is given by*

$$\rho\left(S_1, S_2\right) = \rho\left(N_1, N_2\right) \prod_{j=1}^{2} \mathbb{E}\left[X_j\right] \sqrt{\frac{\mathrm{Var}\left[N_j\right]}{\mathrm{Var}\left[S_j\right]}}, \tag{18}$$

where $\rho\left(N_1, N_2\right)$ *is the correlation coefficient of the bivariate counting distribution of type (10). If, moreover, the claim sizes* $\left(X_{jl}\right)_{l\geq 1}$ *are*

1. *i.i.d., then*

$$\mathbb{E}\left[X_j\right] \sqrt{\frac{Var\left[N_j\right]}{Var\left[S_j\right]}} = \left(\frac{\mathbb{E}\left[N_j\right]}{Var\left[N_j\right]} \frac{Var\left[X_j\right]}{\left(\mathbb{E}\left[X_j\right]\right)^2} + 1\right)^{-1/2};$$

2. *dependent, i.d., with equal covariances* $c_j := \mathrm{cov}\left(X_{ji}, X_{jl}\right)$ *for any* $i \neq l$, $j = 1, 2$, *then*

$$\mathbb{E}\left[X_j\right] \sqrt{\frac{Var\left[N_j\right]}{Var\left[S_j\right]}} = \left(\frac{\mathbb{E}\left[N_j\right]}{Var\left[N_j\right]} \frac{Var\left[X_j\right]}{\left(\mathbb{E}\left[X_j\right]\right)^2} + \frac{c_j \mathbb{E}\left[N_j\left(N_j - 1\right)\right]}{\left(\mathbb{E}\left[X_j\right]\right)^2 Var\left[N_j\right]} + 1\right)^{-1/2}.$$

The following corollary states the fact that the correlation of the compound distribution cannot exceed the one of the counting distribution (its proof is obvious, hence we omit it). Note that when the claims of some type are assumed to be correlated, it is natural to assume that their correlation is positive.

Corollary 4.1. *Under the assumptions of Proposition 4.2, in both cases (i) and (ii) with* $c_j > 0$, $j = 1, 2$, *it holds that* $\rho\left(S_1, S_2\right) < \rho\left(N_1, N_2\right)$.

In consequence, we need the form of $\rho\left(N_1, N_2\right)$. The following result presents some particular formulas useful to the evaluation of $\rho\left(N_1, N_2\right)$ in cases that will be considered later on.

Lemma 4.1. *i. Let N be a discrete r.v. and* $\phi\left(x\right) = e^{-\delta x} - \mathcal{L}_N\left(\delta\right)$. *It holds that*

- *if* $N \sim Po\left(\mu\right)$, *then* $\mathcal{L}_N\left(\delta\right) = e^{\mu\left(e^{-\delta}-1\right)}$ *and* $\frac{\mathbb{E}[N\phi(N)]}{\sqrt{Var[N]}} = \left(e^{-\delta} - 1\right)\sqrt{\mu}e^{\left(e^{-\delta}-1\right)\mu}$;

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

- *if* $N \sim NB(r,q)$, *then* $\mathcal{L}_N(\delta) = \left(\frac{q}{1-(1-q)e^{-\delta}}\right)^r$ *and* $\frac{\mathbb{E}[N\phi(N)]}{\sqrt{Var[N]}} = \frac{(e^{-\delta}-1)q^r\sqrt{r(1-q)}}{(1-(1-q)e^{-\delta})^{r+1}}$;
- *if* $N \sim Log(\theta)$, *then* $\mathcal{L}_N(\delta) = \frac{\ln(1-e^{-\delta}\theta)}{\ln(1-\theta)}$ *and* $\frac{\mathbb{E}[N\phi(N)]}{\sqrt{Var[N]}} = \sqrt{\frac{-\theta}{\theta+\ln(1-\theta)}}\left(\frac{e^{-\delta}(1-\theta)}{1-\theta e^{-\delta}} - \frac{\ln(1-\theta e^{-\delta})}{\ln(1-\theta)}\right)$.

ii. *Assuming now that p denotes the pmf of N and $\phi(x) = p(x) - \sum_n p^2(n)$, hence $E = \sum_{n\in\mathbb{N}} p^2(n)$, we have*

- *if* $N \sim Po(\mu)$, *then* $E = e^{-2\mu}I_0(2\mu)$ *and* $\frac{\mathbb{E}[N\phi(N)]}{\sqrt{Var[N]}} = \sqrt{\mu}e^{-2\mu}(I_1(2\mu) - I_0(2\mu))$;
- *if* $N \sim NB(r,q)$, *then* $E = q^{2r}\ {}_2F_1\left(\{r,r\},\{1\};(1-q)^2\right)$ *and* $\frac{\mathbb{E}[N\phi(N)]}{\sqrt{Var[N]}} = \sqrt{r(1-q)}q^{2r}\Big[rq(1-q)\ {}_2F_1\left(\{r+1,r+1\},\{2\};(1-q)^2\right) - {}_2F_1\left(\{r,r\},\{1\};(1-q)^2\right)\Big]$;
- *if* $N \sim Log(\theta)$, *then* $E = \frac{Li_2(\theta^2)}{\ln^2(1-\theta)}$ *and* $\frac{\mathbb{E}[N\phi(N)]}{\sqrt{Var[N]}} = \frac{(1-\theta)\ln(1-\theta)\ln(1-\theta^2)-\theta Li_2(\theta^2)}{\ln^2(1-\theta)\sqrt{-\theta(\theta+\ln(1-\theta))}}$.

We shall now have a look at some particular choices for the counting and severity distributions.

4.1. Inside-type independency case

In this section, we assume that the claim sizes of type j, i.e., $(X_{jl})_{l\geq 1}$, are independent for any fixed j, and also independent of all the claim sizes of other types, i.e., of $(X_{kl})_{l\geq 1}, \forall k \neq j$. Under these assumptions, Vernic (2017) presented recursive formulas for the evaluation of multivariate compound distributions when the claim sizes distributions are of discrete type. We shall now see how the components of the compound pdf look like in the particular continuous case of independent Erlang distributed claim sizes, i.e., when $X_{jl} \sim \text{Erlang}(k_j, \beta_j)$, $l \geq 1$, $j = \overline{1,m}$. We also assume that each marginal counting distribution, p_j, is of Poisson, negative binomial or logarithmic type. Then, from Proposition 3.1, we know the form of the $p_j \vee h_j$ pdfs, hence, in view of (17), we must find the expressions of the $\tilde{\phi}_j$s. The following results examine the later for the two particular forms of the kernel functions presented in Section 2.2.3.

Proposition 4.3. *Let h be the $Erlang(k,\beta)$ pdf and let $\phi(x) = e^{-\delta x} - \mathcal{L}_N(\delta)$, hence $f(x) = e^{-\delta x}$. It holds the following:*

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

i. *If* $N \sim Po(\mu)$, *then* $((fp) \vee h)(0) = e^{-\mu}$, *while for* $x > 0$,

$$((fp) \vee h)(x) = e^{-\delta-\mu-\beta x} \frac{\mu (\beta x)^k}{x (k-1)!} \times {}_0F_k \left(\{\}, \left\{ \frac{k+1}{k}, \frac{k+2}{k}, \ldots, \frac{2k-1}{k}, 2 \right\}; \frac{\mu}{e^{\delta}} \left(\frac{\beta x}{k} \right)^k \right).$$

ii. *If* $N \sim NB(r, q)$, *then* $((fp) \vee h)(0) = q^r$, *while for* $x > 0$,

$$((fp) \vee h)(x) = \frac{r q^r (1-q) e^{-\delta-\beta x} (\beta x)^k}{x (k-1)!} \times {}_1F_k \left(\{r+1\}, \left\{ \frac{k+1}{k}, \frac{k+2}{k}, \ldots, \frac{2k-1}{k}, 2 \right\}; \frac{1-q}{e^{\delta}} \left(\frac{\beta x}{k} \right)^k \right).$$

iii. *If* $N \sim Log(\theta)$, *then* $((fp) \vee h)(0) = 0$, *while for* $x > 0$,

$$((fp) \vee h)(x) = \begin{cases} -\frac{e^{-\beta x}}{x \ln(1-\theta)} \left(e^{\theta \beta x e^{-\delta}} - 1 \right), k = 1 \\ -\frac{k e^{-\beta x}}{x \ln(1-\theta)} \left[{}_0F_k \left(\{\}, \left\{ \frac{1}{k}, \frac{2}{k}, \ldots, \frac{k-1}{k} \right\}; \frac{\theta}{e^{\delta}} \left(\frac{\beta x}{k} \right)^k \right) - 1 \right], k \geq 2 \end{cases}.$$

The proof of Proposition 4.3 is omitted, being very similar with the proof of Proposition 3.1.

Proposition 4.4. *Let* h *be the* $Erlang(k, \beta)$ *pdf and let* $\phi(x) = p(x) - \sum_{n \in \mathbb{N}} p^2(n)$. *Then*

i. *when* $N \sim Po(\mu)$, *we have* $((fp) \vee h)(0) = e^{-2\mu}$, *while for* $x > 0$,

$$((fp) \vee h)(x) = e^{-2\mu-\beta x} \frac{\mu^2 (\beta x)^k}{x (k-1)!} \times {}_0F_{k+1} \left(\{\}, \left\{ \frac{k+1}{k}, \frac{k+2}{k}, \ldots, \frac{2k-1}{k}, 2, 2 \right\}; \mu^2 \left(\frac{\beta x}{k} \right)^k \right);$$

ii. *when* $N \sim NB(r, q)$, *we have* $((fp) \vee h)(0) = q^{2r}$, *while for* $x > 0$,

$$((fp) \vee h)(x) = \frac{r^2 q^{2r} (1-q)^2 e^{-\beta x} (\beta x)^k}{x (k-1)!} \times {}_2F_{k+1} \left(\{r+1, r+1\}, \left\{ \frac{k+1}{k}, \frac{k+2}{k}, \ldots, \frac{2k-1}{k}, 2, 2 \right\}; (1-q)^2 \left(\frac{\beta x}{k} \right)^k \right);$$

iii. when $N \sim Log(\theta)$, *we have* $((fp) \vee h)(0) = 0$, *while for* $x > 0$,

$$((fp) \vee h)(x)$$
$$= \begin{cases} \dfrac{\theta^2 e^{-\beta x}\beta x}{x \ln^2(1-\theta)}\, {}_2F_2\left(\{1,1\},\{2,2\};\theta^2\beta x\right), & k = 1 \\ \dfrac{\theta^2 e^{-\beta x}(\beta x)^k}{x(k-1)!\ln^2(1-\theta)} & \\ \times {}_2F_{k+1}\left(\{1,1\},\left\{\dfrac{k+1}{k},\dfrac{k+2}{k},\ldots,\dfrac{2k-1}{k},2,2\right\};\theta^2\left(\dfrac{\beta x}{k}\right)^k\right), & k \geq 2 \end{cases}.$$

Example 4.1. *In this example, for illustration purposes, we consider the bivariate case. Hence, for, e.g., two insurance portfolios, we define the compound model (2) as follows:* $(S_1, S_2) = \left(\sum_{l=0}^{N_1} X_{1l}, \sum_{l=0}^{N_2} X_{2l}\right)$ *with* $X_{1l} \sim Erlang(k_1 = 2, \beta_1 = 0.9)$, $X_{2l} \sim Erlang(k_2 = 3, \beta_2 = 0.95), l \geq 1$, $X_{10} = X_{20} = 0$, *while the r.v.s* $N_1 \sim Po(\mu = 2)$, $N_2 \sim NB(r = 4, q = 0.65)$ *are joined by the Sarmanov distribution with kernels of type* $\phi(n) = e^{-n} - \mathcal{L}_N(1)$ *and dependence parameter* $\omega = 3$. *Since in this case we have* $E_1 = \mathcal{L}_{Po(2)}(1) = 0.2825$, $E_2 = \mathcal{L}_{NB(4,0.65)}(1) = 0.3098$, *the pmf of* (N_1, N_2) *is*

$$p(n_1,n_2) = p_1(n_1)\,p_2(n_2)\left[1 + \omega\left(e^{-n_1} - E_1\right)\left(e^{-n_2} - E_2\right)\right]$$
$$= 0.00403\frac{2^{n_1}0.35^{n_2}\Gamma(4+n_2)}{n_1!n_2!}$$
$$\times\left[1 + 3\left(e^{-n_1} - 0.2825\right)\left(e^{-n_2} - 0.3098\right)\right], n_1, n_2 \in \mathbb{N}.$$

Note that this Sarmanov distribution joins different types of marginals, i.e., one Poisson and one negative binomial. Also, its correlation coefficient results using Lemma 4.1 as

$$\rho(N_1,N_2) = \omega\left(e^{-1} - 1\right)^2 \frac{q^r\sqrt{\mu r(1-q)}e^{\left(e^{-1}-1\right)\mu}}{\left(1-(1-q)e^{-1}\right)^{r+1}} \simeq 0.2015.$$

Note that the possible ranges of ω *and* ρ *are in this case* $\omega \in (-2.0192, 4.4983)$, $\rho(N_1, N_2) \in (-0.1356, 0.3021)$.

Thus, according to (17), the joint pdf of (S_1, S_2) *can be written as*

$$(p \vee \mathbf{h})(\mathbf{s}) = \prod_{l=1}^{2}(p_l \vee h_l)(s_l)\left(1 + \omega\tilde{\phi}_1(s_1)\tilde{\phi}_2(s_2)\right), \mathbf{s} \geq \mathbf{0},$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

where $\tilde{\phi}_j = \frac{(f_j p_j)\vee h_j}{p_j \vee h_j} - E_j$ *with* $f_j(x) = e^{-x}$, $j = 1, 2$. *From Proposition 3.1, we obtain the marginal pdfs of* S_1 *and* S_2 *as, respectively,*

$$(p_1 \vee h_1)(s) = \begin{cases} 0.1353, s = 0 \\ 0.2192 s e^{-0.9s}\, {}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 s^2\right), s > 0 \end{cases},$$

$$(p_2 \vee h_2)(s) = \begin{cases} 0.1785, s = 0 \\ 0.1071 s^2 e^{-0.95s}\, {}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 s^3\right), s > 0 \end{cases}.$$

Also, from Proposition 4.3, the kernel functions $\tilde{\phi}_1$ *and* $\tilde{\phi}_2$ *are expressed, respectively, by*

$$\tilde{\phi}_1(0) = 1 - E_1 = 0.7175,$$

$$\tilde{\phi}_1(s) = \frac{{}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 e^{-1} s^2\right)}{{}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 s^2\right) e} - 0.2825, s > 0,$$

$$\tilde{\phi}_2(0) = 1 - E_2 = 0.6902,$$

$$\tilde{\phi}_2(s) = \frac{{}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 e^{-1} s^3\right)}{{}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 s^3\right) e} - 0.3098, s > 0.$$

Then, $(p \vee \mathbf{h})(\mathbf{0}) = p(\mathbf{0}) = 0.06005$, *while for* $s > 0$,

$$\begin{aligned}(p \vee \mathbf{h})(s, 0) &= (p_1 \vee h_1)(s)\, p_2(0) \left(1 + \omega \tilde{\phi}_1(s) \tilde{\phi}_2(0)\right) \\ &= 0.0391 s e^{-0.9s}\, {}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 s^2\right) \\ &\quad \times \left[1 + 2.0706 \left(\frac{{}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 e^{-1} s^2\right)}{{}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 s^2\right) e} - 0.2825\right)\right], \end{aligned} \quad (19)$$

$$\begin{aligned}(p \vee \mathbf{h})(0, s) &= p_1(0)\, (p_2 \vee h_2)(s) \left(1 + \omega \tilde{\phi}_1(0) \tilde{\phi}_2(s)\right) \\ &= 0.0145 s^2 e^{-0.95s}\, {}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 s^3\right) \\ &\quad \times \left[1 + 2.1526 \left(\frac{{}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 e^{-1} s^3\right)}{{}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 s^3\right) e} - 0.3098\right)\right]. \end{aligned} \quad (20)$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

Finally, for $s_1 > 0, s_2 > 0$,

$$(p \vee \mathbf{h})(\mathbf{s}) = 0.0235 s_1 s_2^2 e^{-0.9s_1 - 0.95s_2} \times {}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 s_1^2\right) {}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 s_2^3\right) \times \left[1 + 3\left(\frac{{}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 e^{-1} s_1^2\right)}{{}_0F_2\left(\{\}, \left\{\frac{3}{2}, 2\right\}; 0.405 s_1^2\right) e} - 0.2825\right) \times \left(\frac{{}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 e^{-1} s_2^3\right)}{{}_1F_3\left(\{5\}, \left\{\frac{4}{3}, \frac{5}{3}, 2\right\}; 0.0111 s_2^3\right) e} - 0.3098\right)\right]. \quad (21)$$

In what concerns the correlation of this compound distribution, from Proposition 4.2 item (i), we obtain

$$\rho(S_1, S_2) = \rho(N_1, N_2)\left(\left(\frac{1}{k_1} + 1\right)\left(\frac{q}{k_2} + 1\right)\right)^{-1/2} \simeq 0.1492.$$

Moreover, for the actual values of the parameters, the possible range of this correlation coefficient when ω varies in its interval is $\rho(S_1, S_2) \in (-0.1004, 0.2236)$.

4.2. Inside-type dependency case

We shall now relax the inside-type independency condition while keeping the between-types independency one, i.e., the different types of claims are independent of each other. Hence, we assume that the claim sizes of each type are dependent and follow the multivariate Pareto distribution as defined in Section 2.2.2. As before, we rely on formula (17) to find the compound distribution $p \vee \mathbf{h}$; therefore, we must evaluate all the marginal compound distributions $p_j \vee h_j$, as well as the quantities $(f_j p_j) \vee h_j$ involved in the $\tilde{\phi}_j$s. Considering the previous three particular counting distributions, we already know the form of each $p_j \vee h_j$ from Lemma 3.1, while the expressions of $(f_j p_j) \vee h_j$ are given in the following two properties for two particular kernels cases.

Proposition 4.5. *Let the claim sizes be multivariate type-II Pareto distributed with parameters (α, σ), and let $\phi(x) = e^{-\delta x} - \mathcal{L}_N(\delta)$, hence $f(x) = e^{-\delta x}$. Then,*

i. if $N \sim Po(\mu)$, then $((fp) \vee h)(0) = e^{-\mu}$, while for $x > 0$,

$$((fp) \vee h)(x) = \frac{\alpha \mu e^{-(\mu+\delta)}}{\sigma(1 + x/\sigma)^{1+\alpha}} {}_1F_1\left(\{1+\alpha\}, \{2\}; \frac{\mu x e^{-\delta}}{\sigma + x}\right);$$

ii. if $N \sim NB(r, q)$, then $((fp) \vee h)(0) = q^r$, while for $x > 0$,

$$((fp) \vee h)(x) = \frac{\alpha r q^r (1-q) e^{-\delta}}{\sigma(1 + x/\sigma)^{1+\alpha}} {}_2F_1\left(\{1+\alpha, 1+r\}, \{2\}; \frac{(1-q) x e^{-\delta}}{\sigma + x}\right);$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

iii. if $N \sim Log(\theta)$*, then* $((fp) \vee h)(0) = 0$*, while for* $x > 0$*,*

$$((fp) \vee h)(x) = -\frac{1}{x\ln(1-\theta)}\left[\frac{1}{\left(1+(1-\theta e^{-\delta})x/\sigma\right)^{\alpha}} - \frac{1}{(1+x/\sigma)^{\alpha}}\right].$$

The proof of Proposition 4.5 is similar to the proof of Lemma 3.1 given in Sarabia *et al.* (2016), and therefore is omitted.

Proposition 4.6. *Let the claim sizes be multivariate type-II Pareto distributed with parameters* (α, σ)*, and let* $\phi(x) = p(x) - \sum_{n\in\mathbb{N}} p^2(n)$*, hence* $f = p$ *and* $E = \sum_{n\in\mathbb{N}} p^2(n)$*. Then*

i. when $N \sim Po(\mu)$*, we have* $((fp) \vee h)(0) = e^{-2\mu}$*, while for* $x > 0$*,*

$$((fp) \vee h)(x) = \frac{\alpha e^{-2\mu}\mu^2}{\sigma(1+x/\sigma)^{\alpha+1}}\, {}_1F_2\left(\{1+\alpha\}, \{2,2\}; \frac{\mu^2 x}{\sigma + x}\right);$$

ii. when $N \sim NB(r, q)$*, we have* $((fp) \vee h)(0) = q^{2r}$*, while for* $x > 0$*,*

$$((fp) \vee h)(x) = \frac{\alpha r^2 q^{2r}(1-q)^2}{\sigma(1+x/\sigma)^{\alpha+1}}\, {}_3F_2\left(\{1+\alpha, 1+r, 1+r\}, \{2,2\}; \frac{(1-q)^2 x}{\sigma + x}\right);$$

iii. when $N \sim Log(\theta)$*, we have* $((fp) \vee h)(0) = 0$*, while for* $x > 0$*,*

$$((fp) \vee h)(x) = \frac{\alpha\theta^2}{\sigma(1+x/\sigma)^{\alpha+1}\ln^2(1-\theta)}\, {}_3F_2\left(\{1, 1, 1+\alpha\}, \{2,2\}; \frac{\theta^2 x}{\sigma + x}\right).$$

Example 4.2. *Similarly to Example 4.1, we consider the compound model (2) in the bivariate case, i.e.,* $(S_1, S_2) = \left(\sum_{l=0}^{N_1} X_{1l}, \sum_{l=0}^{N_2} X_{2l}\right)$ *with* $N_1 \sim Po(\mu = 2)$*,* $N_2 \sim Log(\theta = 0.6)$*, while, for* $N_i = n_i, i = 1, 2$*, we let* $(X_{11}, \ldots, X_{1n_1}) \sim PaII_{n_1}(\alpha_1 = 4, \sigma_1 = 3)$ *and* $(X_{21}, \ldots, X_{2n_2}) \sim PaII_{n_2}(\alpha_2 = 3, \sigma_2 = 4)$*. We assume that the r.v.s* N_1 *and* N_2 *are dependent and joined by the Sarmanov distribution with kernels of type* $\phi(n) = e^{-n} - \mathcal{L}_N(1)$ *and dependence parameter* $\omega = 4.5$ *(in this case, the limiting interval is* $\omega \in (-1.9148, 4.8644)$*). Hence, with* $E_1 = \mathcal{L}_{Po(2)}(1) = 0.2825$ *and* $E_2 = \mathcal{L}_{Log(0.6)}(1) = 0.2722$*, the pmf of* (N_1, N_2) *is given by*

$$p(n_1, n_2) = 0.1477\frac{2^{n_1}0.6^{n_2}}{n_1! n_2} \times \left[1 + 4.5\left(e^{-n_1} - 0.2825\right)\left(e^{-n_2} - 0.2722\right)\right], n_1, n_2 \in \mathbb{N}, n_2 > 0.$$

Based on Lemma 4.1, its correlation coefficient is

$$\rho\left(N_1, N_2\right) = \omega\left(e^{-1} - 1\right) e^{\left(e^{-1}-1\right)\mu} \times \sqrt{\frac{-\mu\theta}{\theta + \ln\left(1-\theta\right)}} \left(\frac{1-\theta}{e-\theta} - \frac{\ln\left(1-\theta e^{-1}\right)}{\ln\left(1-\theta\right)}\right) \simeq 0.1304.$$

In fact, depending on ω*, the possible range of this correlation is* $\rho\left(N_1, N_2\right) \in (-0.0555, 0.1410)$.

The joint pdf of (S_1, S_2) *is given by*

$$(p \vee \mathbf{h})(\mathbf{s}) = \prod_{l=1}^{2} (p_l \vee h_l)(s_l)\left(1 + \omega\tilde{\phi}_1(s_1)\,\tilde{\phi}_2(s_2)\right), \mathbf{s} \geq \mathbf{0},$$

with $\tilde{\phi}_j = \frac{(f_j p_j)\vee h_j}{p_j \vee h_j} - E_j$ *and* $f_j(x) = e^{-x}$, $j = 1, 2$. *From Lemma 3.1, we obtain the marginal pdfs of* S_1 *and* S_2*, respectively, as*

$$(p_1 \vee h_1)(s) = \begin{cases} 0.1353, s = 0 \\ 0.3609(1 + s/3)^{-5}\ {}_1F_1\left(\{5\}, \{2\}; 2s/(s+3)\right), s > 0 \end{cases},$$

$$(p_2 \vee h_2)(s) = \begin{cases} 0, s = 0 \\ 1.0913 s^{-1}\left[(1+0.1s)^{-3} - (1+0.25s)^{-3}\right], s > 0 \end{cases}.$$

Also, from Proposition 4.5, the kernel functions $\tilde{\phi}_1$ *and* $\tilde{\phi}_2$ *are expressed, respectively, by*

$$\tilde{\phi}_1(0) = 1 - E_1 = 0.7175,\ \tilde{\phi}_1(s) = \frac{{}_1F_1\left(\{5\}, \{2\}; \frac{2e^{-1}s}{3+s}\right)}{{}_1F_1\left(\{5\}, \{2\}; \frac{2s}{3+s}\right) e} - 0.2825, s > 0,$$

$$\tilde{\phi}_2(0) = 1 - E_2 = 0.7278,$$

$$\tilde{\phi}_2(s) = \frac{(1+0.1948s)^{-3} - (1+0.25s)^{-3}}{(1+0.1s)^{-3} - (1+0.25s)^{-3}} - 0.2722, s > 0.$$

Then, $(p \vee \mathbf{h})(\mathbf{0}) = p(\mathbf{0}) = 0$, *while for* $s > 0$,

$$\begin{aligned}(p \vee \mathbf{h})(s, 0) &= (p_1 \vee h_1)(s)\, p_2(0)\left(1 + \omega\tilde{\phi}_1(s)\,\tilde{\phi}_2(0)\right) \\ &= 0 \text{ as } (p_2 \vee h_2)(0) = 0,\end{aligned}$$

$$\begin{aligned}(p \vee \mathbf{h})(0, s) &= p_1(0)\,(p_2 \vee h_2)(s)\left(1 + \omega\tilde{\phi}_1(0)\,\tilde{\phi}_2(s)\right) \\ &= 0.1477 s^{-1}\left[(1+0.1s)^{-3} - (1+0.25s)^{-3}\right] \\ &\quad \times \left[1 + 3.2289\left(\frac{(1+0.1948s)^{-3} - (1+0.25s)^{-3}}{(1+0.1s)^{-3} - (1+0.25s)^{-3}} - 0.2722\right)\right].\end{aligned}$$

Finally, for $s_1 > 0, s_2 > 0$,

$$(p \vee \mathbf{h})(\mathbf{s}) = \frac{0.3939}{(1+s_1/3)^5 s_2} \, {}_1F_1\left(\{5\}, \{2\}; \frac{2s_1}{s_1+3}\right) \times \left[\frac{1}{(1+0.1s_2)^3} - \frac{1}{(1+0.25s_2)^3}\right] \times \left[1 + 4.5\left(\frac{{}_1F_1\left(\{5\}, \{2\}; \frac{2e^{-1}s_1}{3+s_1}\right)}{{}_1F_1\left(\{5\}, \{2\}; \frac{2s_1}{3+s_1}\right)e} - 0.2825\right) \times \left(\frac{(1+0.1948s_2)^{-3} - (1+0.25s_2)^{-3}}{(1+0.1s_2)^{-3} - (1+0.25s_2)^{-3}} - 0.2722\right)\right].$$

From Proposition 4.2 item (ii), using that $\mathbb{E}[N(N-1)] = Var[N] + \mathbb{E}^2[N] - \mathbb{E}[N]$, *we also find the correlation of the compound distribution as*

$$\rho(S_1, S_2) = \rho(N_1, N_2)\left(\left(\frac{\alpha_1+\mu}{\alpha_1-2}+1\right)\left(\frac{\ln(1-\theta)(\theta+(1-\theta)\alpha_2)}{(\alpha_2-2)(\theta+\ln(1-\theta))}+1\right)\right)^{-\frac{1}{2}} \simeq 0.0262.$$

Considering the actual values of the parameters, when ω *varies in its limiting interval, the possible correlation range is* $\rho(S_1, S_2) \in (-0.0111, 0.0283)$. *Moreover, we numerically maximized (in Mathematica software) the corresponding formula of* $\rho(N_1, N_2)$ *with respect to the parameters* $\mu, \theta, \delta_1, \delta_2$ *and obtained that the maximum possible of the correlation is* $\rho_{max}(N_1, N_2) = 0.4702$ *for* $\mu = 0.2297, \theta = 0.9929, \delta_1 = 1.8367, \delta_2 = 0.0099$, *while* ω *resulted as 6.9073 (note that this time we let the* δ*s vary); hence, when* $\alpha_1 = 4$ *and* $\alpha_2 = 3$, $\rho_{max}(S_1, S_2) = 0.1753$.

5. Numerical example

As mentioned in the Introduction, when dealing with claim size distributions of continuous type, the usual approach consists in discretizing these distributions, and then in evaluating the corresponding discrete compound distribution by applying a specific technique such as, e.g., the recursive method, the Fast Fourier Transform (FFT) algorithm or simulation. Such an approach generates errors starting with the discretization step (by span choice), errors that are usually magnified by applying the specific technique (for some details, see e.g., Robe-Voinea and Vernic, 2016 and the references therein). Therefore, in this example, we compare the cdfs obtained for the bivariate compound distribution presented in Example 4.1 by using the just described approach based on recursions, and by direct calculation.

TABLE 1

COMPARISON OF EXACT AND RECURSIVE CDF VALUES (DIFFERENT SPANS).

CDF	(0, 0)	(0, 5)	(s_1, s_2) (0, 10)	(0, 15)	(0, 20)
Exact	0.006005	0.099282	0.121663	0.13011	0.133381
Rec. $h = 0.001$	0.006005	0.099286	0.121665	0.13011	0.133382
Rec. $h = 0.01$	0.006005	0.099317	0.121676	0.13011	0.133383
Rec. $h = 0.1$	0.006005	0.099634	0.121792	0.13016	0.133399
	(5, 0)	(10, 0)	(15, 0)	(20, 0)	(5, 5)
Exact	0.141177	0.170763	0.177207	0.178323	0.326836
Rec. $h = 0.001$	0.141183	0.170764	0.177207	0.178323	0.326876
Rec. $h = 0.01$	0.141229	0.170774	0.177207	0.178321	0.327223
Rec. $h = 0.1$	0.141554	0.170684	0.177013	0.178106	0.330215
	(10, 10)	(10, 15)	(15, 10)	(15, 15)	(20, 20)
Exact	0.683211	0.812865	0.735079	0.877797	0.955568
Rec. $h = 0.001$	0.683239	0.812886	0.735102	0.877809	0.955573
Rec. $h = 0.01$	0.683479	0.813060	0.735285	0.877902	0.955595
Rec. $h = 0.1$	0.684718	0.813377	0.735876	0.877289	0.954133

In what concerns direct calculation, because the formulas involve the hypergeometric function, we used the facilities provided by the software R and Mathematica to numerically integrate the pdfs (19)–(21). We note that the results were obtained immediately, taking less than a second for each integral.

Regarding the discretization approach, due to the nature of this example, we were able to apply to the resulting discrete compound distribution the recursive method presented in Vernic (2017). The main problem here was the choice of a proper discretization span h; we proceeded by trials (i.e., by successively reducing its value) and, finally, we stopped at $h = 0.001$ (we took the same span for both Erlang marginal distributions).

For comparison, we evaluated the cdf $F(s_1, s_2)$ of (S_1, S_2) by both methods, for $s_1, s_2 \in \{0, 1, 2, \ldots, 20\}$. Some values are presented in Table 1, while in Table 2 we display the maximum absolute error between the exact cdf values and the recursive ones; note how important are the differences between different spans.

On the other hand, the smaller the span is, the longer is the running time and the needed memory space (we wrote both R and Matlab programs). For example, it took more than 1 hour to evaluate the entire discretized cdf for $0 \leq s_1, s_2 \leq 20$ when $h = 0.001$, and we had to optimize the code in order to avoid "out of memory" messages. As another example, to find only the value $F(20, 20)$, the Matlab recursion-discretization code with the span $h = 0.001$ took about 25.14 seconds, while the exact integral value was obtained in about 1 second.

TABLE 2

MAXIMUM ABSOLUTE ERROR BETWEEN THE EXACT AND RECURSIVE CDF VALUES FOR DIFFERENT SPANS.

	$h = 1$	$h = 0.1$	$h = 0.01$	$h = 0.001$
$\max_{s_1, s_2 \in \{0,1,\ldots,20\}} \lvert F - F_{disc.} \rvert (s_1, s_2)$	0.143	0.005	0.00056	0.000045

Therefore, we can see from this example that even if it involves some numerical integrals, direct calculation is more efficient than the classical recursion-discretization method in what concerns the accuracy of the values and the computing time.

6. CONCLUSIONS AND FUTURE WORK

To conclude, in this paper, we obtained some closed-type formulas for the multivariate pdf of some compound distributions with Sarmanov counting distribution and Erlang severities distributions; we also included some dependency between the claim sizes of a certain type by means of a multivariate Pareto distribution. Based on the hypergeometric function which is already implemented in existing software, these formulas seem to be numerically more efficient than the classical recursion-discretization approach, avoiding thus the typical discretization errors generated by the span choice, and the long running time characteristic to the multivariate case.

Therefore, we think that it would be interesting to continue the search for such formulas in the case of compound distributions with continuous severity distributions, formulas expressed by means of special functions already implemented in mathematical software. Moreover, we also plan to pay special attention to the statistical inference of this type of compound distributions.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the two anonymous referees for their valuable comments that helped improving the paper.

REFERENCES

ARNOLD, B.C. (2015) *Pareto Distributions*, 2nd ed. Boca Ratón, FL: Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

ASIMIT, A.V., FURMAN, E. and VERNIC, R. (2010) On a multivariate Pareto distribution. *Insurance Mathematics and Economics*, **46**, 308–316.

ASIMIT, A.V., VERNIC, R. and ZITIKIS, R. (2013) Evaluating risk measures and capital allocations based on multi-losses driven by a heavy-tailed background risk: The multivariate Pareto-II model. *Risks*, **1**(1), 14–33.

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

DENUIT, M., DHAENE, J., GOOVAERTS, M. and KAAS, R. (2005) *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Chichester: John Wiley & Sons.
DUNN, P.K. and SMYTH, G.K. (2005) Series evaluation of Tweedie exponential dispersion model densities. *Statistics and Computing*, **15**(4), 267–280.
GENEST, C., MARCEAU, É. and MESFIOUI, M. (2003) Compound Poisson approximations for individual models with dependent risks. *Insurance: Mathematics and Economics*, **32**(1), 73–91.
GOOVAERTS, M.J. and DHAENE, J. (1996) The compound Poisson approximation for a portfolio of dependent risks. *Insurance: Mathematics and Economics*, **18**(1), 81–85.
GUILLÉN, M., SARABIA, J.M. and PRIETO, F. (2013) Simple risk measure calculations for sums of positive random variables. *Insurance: Mathematics and Economics*, **53**(1), 273–280.
JIN, T. and REN, J. (2014) Recursions and fast Fourier transforms for a new bivariate aggregate claims model. *Scandinavian Actuarial Journal*, **2014**(8), 729–752.
JORGENSEN, B. (1997) *The Theory of Dispersion Models*. London: Chapman and Hall.
KLEIBER, C. and KOTZ, S. (2003) *Statistical Size Distributions in Economics and Actuarial Sciences*, Vol. 470. Hoboken, NJ: John Wiley & Sons.
KLUGMAN, S.A., PANJER, H.H. and WILLMOT, G.E. (1998) *Loss Models: From Data to Decisions*. New York: John Wiley & Sons.
KOTZ, S., BALAKRISHNAN, N. and JOHNSON, N.L. (2000) *Continuous Multivariate Distributions,* Models and Applications, Vol. 1. New York: Wiley.
ROBE-VOINEA, E.G. and VERNIC, R. (2016) On the recursive evaluation of a certain multivariate compound distribution. *Acta Mathematicae Applicatae Sinica, English Series*, **32**(4), 913–920.
SARABIA, J.M., GÓMEZ-DÉNIZ, E., PRIETO, F. and JORDÁ, V. (2016) Risks aggregation in multivariate dependent Pareto distributions. *Insurance: Mathematics and Economics*, **71**, 154–163.
SUNDT, B. and VERNIC, R. (2009) *Recursions for Convolutions and Compound Distributions with Insurance Applications*, EAA Lectures Notes. Berlin: Springer-Verlag.
VERNIC, R. (2017) On the evaluation of some multivariate compound distributions with Sarmanov's counting distribution. SSRN paper 2965053 (submitted).
WILLMOT, G.E. and LIN, X.S. (2011) Risk modelling with the mixed Erlang distribution. *Applied Stochastic Models in Business and Industry*, **27**(1), 2–16.
WILLMOT, G.E. and WOO, J.K. (2015) On some properties of a class of multivariate Erlang mixtures with insurance applications. *Astin Bulletin*, **45**(01), 151–173.

MAISSA TAMRAZ
Department of Actuarial Science
University of Lausanne
UNIL-Dorigny 1015 Lausanne, Switzerland
E-mail: maissa.tamraz@unil.ch

RALUCA VERNIC (Corresponding author)
Faculty of Mathematics and Informatics
Ovidius University of Constanta
124 Mamaia Blvd., 900527 Constanta, Romania
Institute for Mathematical Statistics and Applied Mathematics
Calea 13 Septembrie 13, 050711 Bucharest, Romania
E-mail: rvernic@univ-ovidius.ro

APPENDIX

Proof of Lemma 2.1. We write

$$
\begin{aligned}
((n+1)k-1)!\,(n+1) = k^{nk}&[1\times 2\times \ldots . \times (k-1)]\\
&\times k^{-n}[k\,(2k)\,(3k)\times \ldots . \times (nk)]\,(n+1)\\
&\times k^{-n}[(k+1)\,(2k+1)\times \ldots . \times (nk+1)]\\
&\times k^{-n}[(k+2)\,(2k+2)\times \ldots . \times (nk+2)]\times \ldots . \times\\
&\times k^{-n}[(k+(k-1))\,(2k+(k-1))\times \ldots . \times (nk+(k-1))]
\end{aligned}
$$

$$
\begin{aligned}
&= (k-1)!k^{nk}\,(2\times(2+1)\times \ldots . \times (2+n-1))\\
&\quad\times\left[\frac{k+1}{k}\left(\frac{k+1}{k}+1\right)\times \ldots . \times\left(\frac{k+1}{k}+n-1\right)\right]\\
&\quad\times \ldots \times\left[\frac{k+(k-1)}{k}\left(\frac{k+(k-1)}{k}+1\right)\times \ldots . \times\left(\frac{k+(k-1)}{k}+n-1\right)\right],
\end{aligned}
$$

from where formula (8) is immediate. To obtain formula (9), we similarly write

$$
\begin{aligned}
(nk)! &= [(k+1)\,(2k+1)\times \ldots . \times ((n-1)\,k+1)]\\
&\quad\times[2\,(k+2)\,(2k+2)\times \ldots . \times ((n-1)\,k+2)]\times \ldots . \times\\
&\quad\times[(k-1)\,(k+(k-1))\,(2k+(k-1))\times \ldots . \times ((n-1)\,k+(k-1))]\\
&\quad\times[k\,(2k)\,(3k)\times \ldots . \times (nk)]\\
&= k^{nk}\left[\frac{1}{k}\left(\frac{1}{k}+1\right)\times \ldots . \times\left(\frac{1}{k}+n-1\right)\right]\\
&\quad\times\left[\frac{2}{k}\left(\frac{2}{k}+1\right)\times \ldots . \times\left(\frac{2}{k}+n-1\right)\right]\times \ldots . \times\\
&\quad\times\left[\frac{k-1}{k}\left(\frac{k-1}{k}+1\right)\times \ldots . \times\left(\frac{k-1}{k}+n-1\right)\right]n!,
\end{aligned}
$$

hence the result. ■

Proof of Proposition 3.1. (i) In the Poisson case, based on (15) and (8), we have for $x > 0$,

$$
\begin{aligned}
(p\vee h)\,(x) &= \sum_{n=1}^{\infty} e^{-\mu}\frac{\mu^n}{n!}\frac{\beta^{nk}}{(nk-1)!}x^{nk-1}e^{-\beta x} = e^{-\mu-\beta x}\sum_{n=1}^{\infty}\frac{\left(\mu\,(\beta x)^k\right)^n}{n!\,(nk-1)!}x^{-1}\\
&= e^{-\mu-\beta x}\frac{\mu\,(\beta x)^k}{x}\sum_{n=0}^{\infty}\frac{\left(\mu\,(\beta x)^k\right)^n}{(n+1)!\,((n+1)\,k-1)!}\\
&= e^{-\mu-\beta x}\frac{\mu\,(\beta x)^k}{x\,(k-1)!}\sum_{n=0}^{\infty}\frac{\left(\mu\left(\frac{\beta x}{k}\right)^k\right)^n}{n!}\frac{1}{\left(\frac{k+1}{k}\right)_{(n)}\left(\frac{k+2}{k}\right)_{(n)}\times\ldots\times\left(\frac{k+(k-1)}{k}\right)_{(n)}(2)_{(n)}},
\end{aligned}
$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

and using the definition (7), the formula is immediate. In the negative binomial case (*ii*), (15) and (8) yields

$$(p \vee h)(x) = \frac{q^r e^{-\beta x}}{\Gamma(r)} \sum_{n=1}^{\infty} \frac{\Gamma(r+n)}{n!} (1-q)^n \frac{\beta^{nk}}{(nk-1)!} x^{nk-1}$$

$$= \frac{q^r e^{-\beta x}}{x\Gamma(r)} \sum_{n=1}^{\infty} \frac{\left((1-q)(\beta x)^k\right)^n}{n!} \frac{\Gamma(r+n)}{(nk-1)!}$$

$$= \frac{q^r e^{-\beta x}}{x} \sum_{n=0}^{\infty} \frac{\left((1-q)(\beta x)^k\right)^{n+1}}{n!} \frac{\Gamma(r+n+1)}{(n+1)((n+1)k-1)!\Gamma(r)}$$

$$= \frac{q^r(1-q)(\beta x)^k e^{-\beta x}}{x} \sum_{n=0}^{\infty} \frac{\left((1-q)(\beta x)^k\right)^n}{n!(k-1)!k^{nk}}$$

$$\times \frac{r(r+1)\times\ldots\times(r+1+n-1)}{\left(\frac{k+1}{k}\right)_{(n)}\left(\frac{k+2}{k}\right)_{(n)}\times\ldots\times\left(\frac{k+(k-1)}{k}\right)_{(n)}(2)_{(n)}},$$

hence the result. Finally, for the logarithmic distribution, (15) gives

$$(p \vee h)(x) = -\frac{e^{-\beta x}}{\ln(1-\theta)} \sum_{n=1}^{\infty} \frac{\theta^n}{n} \frac{\beta^{nk} x^{nk-1}}{(nk-1)!} = -\frac{e^{-\beta x} k}{x\ln(1-\theta)} \sum_{n=1}^{\infty} \frac{\left(\theta(\beta x)^k\right)^n}{(nk)!}$$

$$= -\frac{ke^{-\beta x}}{x\ln(1-\theta)} \sum_{n=0}^{\infty} \left[\frac{\left(\theta(\beta x)^k\right)^n}{(nk)!} - 1\right].$$

When $k = 1$ (i.e., exponentially distributed claims), we obtain

$$(p \vee h)(x) = -\frac{e^{-\beta x}}{x\ln(1-\theta)} \left(e^{\theta\beta x} - 1\right),$$

otherwise, using (9), we have

$$(p \vee h)(x) = -\frac{ke^{-\beta x}}{x\ln(1-\theta)} \sum_{n=0}^{\infty} \left[\frac{\left(\theta(\beta x)^k\right)^n}{n!k^{nk}} \frac{1}{\left(\frac{1}{k}\right)_{(n)}\left(\frac{2}{k}\right)_{(n)}\cdots\left(\frac{k-1}{k}\right)_{(n)}} - 1\right],$$

i.e., the second formula of (*iii*), which completes the proof. ∎

Proof of Proposition 4.1. This result was proved by Vernic (2017) in the case when both inside-type and between-types independency assumptions hold and the claim sizes are of discrete type. Considering now the case with only between-types independence and no inside-type independence, along with continuous claim sizes, the proof is similar: inserting formula

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

(10) and $\phi_l(x) = f_l(x) - E_l$ into (6), we have

$$\begin{aligned}
(p \vee \mathbf{h})(\mathbf{s}) &= \sum_{n_1=0}^{\infty} \ldots \sum_{n_m=0}^{\infty} \left(\prod_{l=1}^{m} p_l(n_l)\, h_l^{(n_l)}(s_l) \right) \\
&\quad \times \left(1 + \sum_{k=2}^{m} \sum_{1 \le j_1 < \ldots < j_k \le m} \omega_{j_1 \ldots j_k}\, \phi_{j_1}(n_{j_1}) \times \ldots \times \phi_{j_k}(n_{j_k}) \right) \\
&= \prod_{l=1}^{m} \left(\sum_{n_l=0}^{\infty} p_l(n_l)\, h_l^{(n_l)}(s_l) \right) + \sum_{k=2}^{m} \sum_{1 \le j_1 < \ldots < j_k \le m} \omega_{j_1 \ldots j_k} \\
&\quad \times \prod_{l=1}^{k} \left(\sum_{n_{j_l}=0}^{\infty} \left(f_{j_l}(n_{j_l}) - E_{j_l} \right) p_{j_l}(n_{j_l})\, h_{j_l}^{(n_{j_l})}(s_{j_l}) \right) \\
&\quad \times \prod_{l=k+1}^{m} \left(\sum_{n_{j_l}=0}^{\infty} p_{j_l}(n_{j_l})\, h_{j_l}^{(n_{j_l})}(s_{j_l}) \right),
\end{aligned}$$

from where, using (3), we obtain formula (16). This formula easily yields the Sarmanov form (17); to verify that $\tilde{\phi}_j$ is indeed in the form (13), we proceed as follows: we denote by S_j the marginal r.v. having the compound distribution $p_j \vee h_j$, hence

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{(f_j p_j) \vee h_j}{p_j \vee h_j} \right)(S_j) \right] &= \int_0^{\infty} \left(\frac{(f_j p_j) \vee h_j}{p_j \vee h_j} \right)(s) \left(p_j \vee h_j \right)(s)\, ds \\
&= \int_0^{\infty} \left((f_j p_j) \vee h_j \right)(s)\, ds \\
&= \int_0^{\infty} \sum_{k=0}^{\infty} (f_j p_j)(k)\, h_j^{(k)}(s)\, ds = \sum_{k=0}^{\infty} (f_j p_j)(k) \int_0^{\infty} h_j^{(k)}(s)\, ds \\
&= \sum_{k=0}^{\infty} (f_j p_j)(k) = \mathbb{E}\left[f_j(N_j) \right] = E_j,
\end{aligned}$$

which completes the proof. ∎

Proof of Proposition 4.2. To find $\rho(S_1, S_2)$, according to formula (14), we must evaluate

$$\begin{aligned}
\mathbb{E}\left[S_j \tilde{\phi}_j(S_j) \right] &= \int_0^{\infty} x \left(\frac{(f_j p_j) \vee h_j}{p_j \vee h_j} \right)(x) \left(p_j \vee h_j \right)(x)\, dx - E_j \mathbb{E}\left[S_j \right] \\
&= \sum_{n=0}^{\infty} (f_j p_j)(n) \int_0^{\infty} x h_j^{(n)}(x)\, dx - E_j \mathbb{E}\left[S_j \right].
\end{aligned}$$

In both cases with independent and dependent claim sizes, it holds that $\mathbb{E}\left[S_j\right] = \mathbb{E}\left[N_j\right] \mathbb{E}\left[X_j\right]$ (for the dependent case, see Sarabia *et al.*, 2016); also, since $h_j^{(n)}$ is the pdf

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

of $X_{j1}+\cdots+X_{jn}$, we have $\int_0^\infty x h_j^{(n)}(x)\,dx = n\mathbb{E}\left[X_j\right]$, hence

$$\mathbb{E}\left[S_j\tilde{\phi}_j\left(S_j\right)\right] = \mathbb{E}\left[X_j\right]\sum_{n=0}^{\infty} n\left(f_j p_j\right)(n) - E_j\mathbb{E}\left[N_j\right]\mathbb{E}\left[X_j\right]$$
$$= \mathbb{E}\left[X_j\right]\left(\mathbb{E}\left[N_j f_j\left(N_j\right)\right] - E_j\mathbb{E}\left[N_j\right]\right) = \mathbb{E}\left[X_j\right]\mathbb{E}\left[N_j\phi_j\left(N_j\right)\right].$$

Therefore,

$$\rho\left(S_1, S_2\right) = \omega\prod_{j=1}^{2}\frac{\mathbb{E}\left[S_j\tilde{\phi}_j\left(S_j\right)\right]}{\sqrt{Var\left[S_j\right]}} = \rho\left(N_1, N_2\right)\prod_{j=1}^{2}\mathbb{E}\left[X_j\right]\sqrt{\frac{Var\left[N_j\right]}{Var\left[S_j\right]}}.$$

i. Under the independence assumption, we have $Var\left[S_j\right] = \mathbb{E}\left[N_j\right]Var\left[X_j\right] + \mathbb{E}^2\left[X_j\right]Var\left[N_j\right]$, which inserted into (18) easily yields the corresponding formula.

ii. Under this dependence assumption, Sarabia *et al.* (2016) proved that

$$Var\left[S_j\right] = \mathbb{E}\left[N_j\right]Var\left[X_j\right] + \mathbb{E}^2\left[X_j\right]Var\left[N_j\right] + c_j\mathbb{E}\left[N_j\left(N_j-1\right)\right],$$

and inserting it into (18) we obtain the last formula. This completes the proof. ∎

Proof of Lemma 4.1. i. When $\phi(x) = e^{-\delta x} - \mathcal{L}_N(\delta)$, we have

$$\mathbb{E}\left[N\phi(N)\right] = \mathbb{E}\left[Ne^{-\delta N}\right] - \mathcal{L}_N(\delta)\,\mathbb{E}[N].$$

For $N\sim \text{Po}(\mu)$, $\mathcal{L}_N(\delta) = e^{\mu\left(e^{-\delta}-1\right)}$, while

$$\mathbb{E}\left[N\phi(N)\right] = \sum_{n=1}^{\infty} ne^{-\delta n}e^{-\mu}\frac{\mu^n}{n!} - \mu e^{\mu\left(e^{-\delta}-1\right)} = \mu e^{-\mu-\delta}\sum_{n=0}^{\infty}\frac{\left(\mu e^{-\delta}\right)^n}{n!} - \mu e^{\mu\left(e^{-\delta}-1\right)}$$
$$= \mu\left(e^{\mu\left(e^{-\delta}-1\right)-\delta} - e^{\mu\left(e^{-\delta}-1\right)}\right),$$

from where we immediately obtain the stated formula. For $N\sim \text{NB}(r,q)$, $\mathcal{L}_N(\delta) = \frac{q^r}{\left(1-(1-q)e^{-\delta}\right)^r}$ and

$$\mathbb{E}\left[Ne^{-\delta N}\right] = \sum_{n=0}^{\infty} ne^{-\delta n}\frac{\Gamma(r+n)}{n!\Gamma(r)}q^r(1-q)^n = q^r\sum_{n=0}^{\infty} n\frac{\Gamma(r+n)}{n!\Gamma(r)}\left((1-q)e^{-\delta}\right)^n$$
$$= \frac{q^r}{\left(1-(1-q)e^{-\delta}\right)^r}\frac{r(1-q)e^{-\delta}}{1-(1-q)e^{-\delta}} = \frac{r(1-q)q^re^{-\delta}}{\left(1-(1-q)e^{-\delta}\right)^{r+1}},$$

hence,

$$\mathbb{E}\left[N\phi(N)\right] = \frac{r(1-q)q^re^{-\delta}}{\left(1-(1-q)e^{-\delta}\right)^{r+1}} - \frac{q^{r-1}r(1-q)}{\left(1-(1-q)e^{-\delta}\right)^r}$$
$$= \frac{r(1-q)q^{r-1}}{\left(1-(1-q)e^{-\delta}\right)^r}\left(\frac{qe^{-\delta}}{1-(1-q)e^{-\delta}} - 1\right) = \frac{\left(e^{-\delta}-1\right)r(1-q)q^{r-1}}{\left(1-(1-q)e^{-\delta}\right)^{r+1}},$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

which easily yields the stated formula. When $N \sim \text{Log}(\theta)$, $\mathcal{L}_N(\delta) = \frac{\ln(1-\theta e^{-\delta})}{\ln(1-\theta)}$ and using

$$\mathbb{E}\left[Ne^{-\delta N}\right] = -\frac{1}{\ln(1-\theta)}\sum_{n=1}^{\infty} ne^{-\delta n}\frac{\theta^n}{n} = -\frac{\theta e^{-\delta}}{\ln(1-\theta)}\sum_{n=0}^{\infty}\left(e^{-\delta}\theta\right)^n$$
$$= -\frac{\theta e^{-\delta}}{\ln(1-\theta)}\frac{1}{1-e^{-\delta}\theta},$$

we have

$$\frac{\mathbb{E}[N\phi(N)]}{\sqrt{Var[N]}} = \left(\frac{-\theta e^{-\delta}}{\left(1-\theta e^{-\delta}\right)\ln(1-\theta)} + \frac{\theta\ln\left(1-\theta e^{-\delta}\right)}{(1-\theta)\ln^2(1-\theta)}\right)\sqrt{\frac{(1-\theta)^2\ln^2(1-\theta)}{-\theta(\theta+\ln(1-\theta))}}$$
$$= \frac{-(1-\theta)\sqrt{\theta}}{\sqrt{-(\theta+\ln(1-\theta))}}\left(\frac{-e^{-\delta}}{1-\theta e^{-\delta}} + \frac{\ln\left(1-\theta e^{-\delta}\right)}{(1-\theta)\ln(1-\theta)}\right),$$

from where results the last formula of case (i).

ii. Let now $\phi(x) = p(x) - \sum_n p^2(n)$. If $N \sim \text{Po}(\mu)$, we have

$$E = \sum_{n\in\mathbb{N}} p^2(n) = \sum_{n\in\mathbb{N}} e^{-2\mu}\frac{\mu^{2n}}{(n!)^2} = e^{-2\mu}I_0(2\mu),$$

while

$$\mathbb{E}[N\phi(N)] = \sum_{n=1}^{\infty} n\left(e^{-\mu}\frac{\mu^n}{n!}\right)^2 - \mu E = \mu e^{-2\mu}\sum_{n=0}^{\infty}\frac{\mu^{2n+1}}{n!(n+1)!} - \mu e^{-2\mu}I_0(2\mu)$$
$$= \mu e^{-2\mu}\left(I_1(2\mu) - I_0(2\mu)\right),$$

which immediately yields the corresponding result. When $N \sim \text{NB}(r, q)$, we calculate

$$E = \sum_{n\in\mathbb{N}}\left[\frac{\Gamma(r+n)}{\Gamma(r)}\right]^2\frac{q^{2r}(1-q)^{2n}}{(n!)^2} = q^{2r}\sum_{n=0}^{\infty}\frac{(1-q)^{2n}}{n!}\frac{[r(r+1)\times\ldots\times(r+n-1)]^2}{1\times 2\times\ldots\times(1+n-1)}$$
$$= q^{2r}\,{}_2F_1\left(\{r, r\}, \{1\}; (1-q)^2\right),$$

and

$$\mathbb{E}[Np(N)] = q^{2r}\sum_{n=1}^{\infty} n(1-q)^{2n}\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\right)^2$$
$$= q^{2r}(1-q)^2\sum_{n=0}^{\infty}\frac{(1-q)^{2n}}{n!(n+1)!}\left(\frac{\Gamma(r+n+1)}{\Gamma(r)}\right)^2$$
$$= r^2q^{2r}(1-q)^2\sum_{n=0}^{\infty}\frac{(1-q)^{2n}}{n!}\frac{[(r+1)\times\ldots\times(r+1+n-1)]^2}{2\times\ldots\times(n+2-1)}$$
$$= r^2q^{2r}(1-q)^2\,{}_2F_1\left(\{r+1, r+1\}, \{2\}; (1-q)^2\right).$$

Thus, based on the above, a straightforward calculation yields the stated formula. For $N \sim \text{Log}(\theta)$, we have

$$E = \sum_{n=1}^{\infty} p^2(n) = \frac{1}{\ln^2(1-\theta)} \sum_{n=1}^{\infty} \frac{\theta^{2n}}{n^2} = \frac{1}{\ln^2(1-\theta)} Li_2\left(\theta^2\right),$$

and

$$\mathbb{E}[Np(N)] = \sum_{n=1}^{\infty} \frac{n}{\ln^2(1-\theta)} \frac{\theta^{2n}}{n^2} = \frac{1}{\ln^2(1-\theta)} \sum_{n=1}^{\infty} \frac{\theta^{2n}}{n} = -\frac{\ln(1-\theta^2)}{\ln^2(1-\theta)}.$$

It follows that

$$\frac{\mathbb{E}[N\phi(N)]}{\sqrt{Var[N]}} = \left(\frac{\theta Li_2\left(\theta^2\right)}{(1-\theta)\ln^3(1-\theta)} - \frac{\ln(1-\theta^2)}{\ln^2(1-\theta)}\right)\sqrt{\frac{(1-\theta)^2\ln^2(1-\theta)}{-\theta(\theta+\ln(1-\theta))}},$$

which easily completes the proof. ■

Proof of Proposition 4.4. (*i*) For $N \sim \text{Po}(\mu)$, we have $((fp) \vee h)(0) = (fp)(0) = p^2(0) = e^{-2\mu}$, while for $x > 0$, using (8) and $(n+1)! = (2)_{(n)}$,

$$\begin{aligned}
((fp) \vee h)(x) &= \sum_{n=1}^{\infty} e^{-2\mu} \frac{\mu^{2n}}{(n!)^2} \frac{\beta^{nk}}{(nk-1)!} x^{nk-1} e^{-\beta x} \\
&= e^{-2\mu-\beta x} \frac{\mu^2 (\beta x)^k}{x} \sum_{n=0}^{\infty} \frac{\left(\mu^2 (\beta x)^k\right)^n}{[(n+1)!]^2 ((n+1)k-1)!} \\
&= e^{-2\mu-\beta x} \frac{\mu^2 (\beta x)^k}{x(k-1)!} \sum_{n=0}^{\infty} \frac{\left(\mu^2 (\beta x/k)^k\right)^n}{n!} \\
&\quad \times \frac{1}{\left(\frac{k+1}{k}\right)_{(n)} \left(\frac{k+2}{k}\right)_{(n)} \times \ldots \times \left(\frac{k+(k-1)}{k}\right)_{(n)} (2)_{(n)} (2)_{(n)}},
\end{aligned}$$

hence the stated formula. In the case (*ii*) when $N \sim \text{NB}(r, q)$, we obtain $((fp) \vee h)(0) = p^2(0) = q^{2r}$, and for $x > 0$, using (8),

$$\begin{aligned}
((fp) \vee h)(x) &= q^{2r} e^{-\beta x} \sum_{n=1}^{\infty} \left[\frac{\Gamma(r+n)}{\Gamma(r)}\right]^2 \frac{(1-q)^{2n}}{(n!)^2} \frac{\beta^{nk} x^{nk-1}}{(nk-1)!} \\
&= \frac{q^{2r} e^{-\beta x} (1-q)^2 (\beta x)^k}{x} \sum_{n=0}^{\infty} \frac{\left((1-q)^2 (\beta x)^k\right)^n}{n!} \\
&\quad \times \frac{[r(r+1) \times \ldots \times (r+1+n-1)]^2}{(n+1)(n+1)!((n+1)k-1)!} \\
&= \frac{r^2 q^{2r} e^{-\beta x} (1-q)^2 (\beta x)^k}{x(k-1)!} \sum_{n=0}^{\infty} \frac{\left((1-q)^2 (\beta x)^k\right)^n}{n! k^{nk}} \\
&\quad \times \frac{(r+1)_{(n)} (r+1)_{(n)}}{\left(\frac{k+1}{k}\right)_{(n)} \left(\frac{k+2}{k}\right)_{(n)} \times \ldots \times \left(\frac{k+(k-1)}{k}\right)_{(n)} (2)_{(n)} (2)_{(n)}},
\end{aligned}$$

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press

which immediately yields the result. To prove the formulas in case (*iii*) when $N \sim \text{Log}(\theta)$ and clearly $((fp) \vee h)(0) = 0$, for $x > 0$, we use

$$((fp) \vee h)(x) = \frac{e^{-\beta x}}{\ln^2(1-\theta)} \sum_{n=1}^{\infty} \frac{\theta^{2n}}{n^2} \frac{\beta^{nk} x^{nk-1}}{(nk-1)!} = \frac{e^{-\beta x}\theta^2 (\beta x)^k}{x \ln^2(1-\theta)} \sum_{n=0}^{\infty} \frac{\left(\theta^2 (\beta x)^k\right)^n}{(n+1)^2 ((n+1)k-1)!}$$
$$= \frac{e^{-\beta x}\theta^2 (\beta x)^k}{x \ln^2(1-\theta)} \sum_{n=0}^{\infty} \frac{\left(\theta^2 (\beta x)^k\right)^n}{n!} \frac{n!}{(n+1)((n+1)k-1)!} \frac{n!}{(n+1)!}.$$

When $k = 1$, this gives

$$((fp) \vee h)(x) = \frac{e^{-\beta x}\theta^2 \beta x}{x \ln^2(1-\theta)} \sum_{n=0}^{\infty} \frac{\left(\theta^2 \beta x\right)^n}{n!} \frac{n!}{(n+1)!} \frac{n!}{(n+1)!}$$
$$= \frac{\theta^2 e^{-\beta x} \beta x}{x \ln^2(1-\theta)} \, {}_2F_2\left(\{1,1\}, \{2,2\}; \theta^2 \beta x\right),$$

while for $k \geq 2$, we apply formula (8) and obtain the result. ■

Proof of Proposition 4.6. i. For $N \sim \text{Po}(\mu)$, we get

$$((fp) \vee h)(x) = \sum_{n=1}^{\infty} e^{-2\mu} \frac{\mu^{2n}}{(n!)^2} \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\Gamma(n)} \frac{x^{n-1}}{\sigma^n (1+x/\sigma)^{n+\alpha}}$$
$$= \frac{e^{-2\mu}}{x(1+x/\sigma)^{\alpha}} \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+1+n)}{\Gamma(\alpha)\Gamma(n+1)} \frac{\mu^{2(n+1)}}{[(n+1)!]^2} \left(\frac{x}{\sigma(1+x/\sigma)}\right)^{n+1}$$
$$= \frac{\alpha \mu^2 e^{-2\mu}}{\sigma(1+x/\sigma)^{1+\alpha}} \sum_{n=0}^{\infty} \left(\frac{\mu^2 x}{\sigma + x}\right)^n \frac{1}{n!} \frac{(1+\alpha)_{(n)}}{(2)_{(n)}(2)_n},$$

hence the result.

ii. For $N \sim \text{NB}(r, q)$, we have

$$((fp) \vee h)(x)$$
$$= \sum_{n=1}^{\infty} \left[\frac{\Gamma(n+r)}{\Gamma(r) n!} q^r (1-q)^n\right]^2 \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\Gamma(n)} \frac{x^{n-1}}{\sigma^n (1+x/\sigma)^{n+\alpha}}$$
$$= \frac{q^{2r}}{x\left(1+\frac{x}{\sigma}\right)^{\alpha}} \sum_{n=0}^{\infty} \frac{(1-q)^{2n+2}}{n!} \left(\frac{x}{\sigma\left(1+\frac{x}{\sigma}\right)}\right)^{n+1} \left[\frac{\Gamma(r+1+n)}{\Gamma(r)(n+1)!}\right]^2 \frac{\Gamma(\alpha+1+n)}{\Gamma(\alpha)}$$
$$= \frac{\alpha r^2 q^{2r} (1-q)^2}{\sigma(1+x/\sigma)^{\alpha+1}} \sum_{n=0}^{\infty} \left(\frac{(1-q)^2 x}{\sigma + x}\right)^n \frac{1}{n!} \frac{(1+\alpha)_{(n)}(1+r)_{(n)}(1+r)_{(n)}}{(2)_{(n)}(2)_{(n)}},$$

yielding the result. In the case (*iii*) where $N \sim \mathrm{Log}(\theta)$, we write

$$\begin{aligned}((fp) \vee h)(x) &= \left[\frac{1}{\ln(1-\theta)}\right]^2 \sum_{n=1}^{\infty} \left(\frac{\theta^n}{n}\right)^2 \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\Gamma(n)} \frac{x^{n-1}}{\sigma^n (1+x/\sigma)^{n+\alpha}} \\ &= \frac{1}{\ln^2(1-\theta)} \frac{1}{x(1+x/\sigma)^{\alpha}} \sum_{n=0}^{\infty} \frac{\theta^{2n+2}}{(n+1)^2} \frac{\Gamma(\alpha+n+1)}{\Gamma(\alpha)\Gamma(n+1)} \left(\frac{x}{\sigma(1+x/\sigma)}\right)^{n+1} \\ &= \frac{1}{\ln^2(1-\theta)} \frac{\alpha\theta^2}{\sigma(1+x/\sigma)^{\alpha+1}} \sum_{n=0}^{\infty} \left(\frac{\theta^2 x}{\sigma+x}\right)^n \frac{1}{n!} \frac{(\alpha+1)_{(n)}}{(n+1)^2},\end{aligned}$$

which, using $\frac{1}{n+1} = \frac{n!}{(n+1)!}$, leads to the stated formula. ■

https://doi.org/10.1017/asb.2017.46 Published online by Cambridge University Press