

Article

Frequency and Severity Dependence in the Collective Risk Model: An Approach Based on Sarmanov Distribution

Catalina Bolancé ^{1,†}  and Raluca Vernic ^{2,*,†} 

¹ Department of Econometrics, Riskcenter-IREA University of Barcelona, Av. Diagonal, 690, 08034 Barcelona, Spain; bolance@ub.edu

² Faculty of Mathematics and Computer Science, Ovidius University of Constanta, 124 Mamaia Blvd., Constanta, and Gheorghe Mihoc-Caius Iacob Institute of Mathematical Statistics and Applied Mathematics of the Romanian Academy, Calea 13 Septembrie 13, 050711 Bucharest, Romania

* Correspondence: rvernic@univ-ovidius.ro

† These authors contributed equally to this work.

Received: 22 July 2020; Accepted: 18 August 2020; Published: 21 August 2020

Abstract: In actuarial mathematics, the claims of an insurance portfolio are often modeled using the collective risk model, which consists of a random number of claims of independent, identically distributed (i.i.d.) random variables (r.v.s) that represent cost per claim. To facilitate computations, there is a classical assumption of independence between the random number of such random variables (i.e., the claims frequency) and the random variables themselves (i.e., the claim severities). However, recent studies showed that, in practice, this assumption does not always hold, hence, introducing dependence in the collective model becomes a necessity. In this sense, one trend consists of assuming dependence between the number of claims and their average severity. Alternatively, we can consider heterogeneity between the individual cost of claims associated with a given number of claims. Using the Sarmanov distribution, in this paper we aim at introducing dependence between the number of claims and the individual claim severities. As marginal models, we use the Poisson and Negative Binomial (NB) distributions for the number of claims, and the Gamma and Lognormal distributions for the cost of claims. The maximum likelihood estimation of the proposed Sarmanov distribution is discussed. We present a numerical study using a real data set from a Spanish insurance portfolio.

Keywords: collective risk model; dependence; bivariate Sarmanov distribution; estimation

MSC: 60E05; 62H05; 62H10

1. Introduction

The collective risk model is a basic classical actuarial risk model consisting of the sum of a random number of independent, identically distributed (i.i.d.) random variables (r.v.s) that represent costs. To facilitate computations related to this model, there is a classical assumption of independence between the random number of such random variables (i.e., the claim frequency) and the random variables themselves (i.e., the claim severities). However, studies on real data emphasized in several cases the existence of a certain dependence that should be taken into account because it can affect important actuarial quantities like premiums and ruin probabilities. Therefore, alternative approaches incorporate dependence between the number of claims and their average severity, see, for example, Erhardt and Czado [1], Czado et al. [2], Krämer et al. [3], Lee and Shi [4], or Oh et al. [5]. In this paper, we propose the bivariate Sarmanov distribution to analyze the joint behavior of the number of claims and of each one of the individual claim amounts, instead of their average; i.e., we consider heterogeneity

between the claim amounts associated with each number of claims. Therefore, we need to work with all the individual claim amounts separately, and some analytical results become complex, like e.g., the distribution of the aggregated claims and its moments, which, in insurance, are fundamental to calculate the premium.

Starting from an increasing need of flexible multivariate distributions, Sarmanov distribution recently gained interest in the actuarial literature and, fitted to some real insurance data in its bivariate and trivariate forms, provided better fits than other distributions, including Copula ones. In this sense, we mention its applications in modeling continuous claim sizes see [6], modeling discrete claim frequencies see [7,8], in the evaluation of ruin probabilities see [9,10] or in capital allocation see [11–13].

Therefore, in this work, we use the Sarmanov distribution to model the joint distribution of the frequency and of the individual severity of claims, and we deduce the moments of the distribution of the resulting aggregate claims in the collective model. In Section 2, we present general theoretical results. In Section 3, we study some interesting particular distributions that are commonly used in insurance, i.e., the Poisson and Negative Binomial (NB) distribution for frequency and the Gamma and Lognormal for severity. A numerical example is presented in Section 4: on a real data set, we compare the bivariate Sarmanov distribution relating the frequency and the individual severity with the simpler case in which it is assumed that the claim amounts are all equal to the mean cost per policyholder; this last assumption implies eliminating the heterogeneity within each insured, and has been used in the alternative works cited in the first paragraph of this introduction. Finally, we conclude in Section 5. All the proofs are given in the Appendix A.

2. Collective Model with Frequency Dependent on the Individual Claims

2.1. Introducing Sarmanov Dependence

If N denotes the r.v. number of claims from a certain portfolio and X_1, X_2, \dots, X_N the corresponding claim amounts, then the resulting aggregate claims can be represented by the collective model as

$$S = \sum_{j=1}^N X_j, \tag{1}$$

where $S = 0$ when $N = 0$. The usual assumptions under which this model is considered are X_1, X_2, \dots, X_N i.i.d. positive r.v.s, independent of the r.v. N . If, in particular, we assume that $X_1 = X_2 = \dots = X_N = \bar{X}$, where \bar{X} is the mean cost per policyholder, we obtain a simpler representation of the collective risk model, for which $S = N\bar{X}$.

Let X denote the generic r.v. claim amount whose distribution is assumed to be absolutely continuous with probability density function (pdf) denoted by f_X , let p denote the probability mass function (pmf) of N and let f_S be pdf of S . The cumulative distribution function (cdf) of an r.v. will be denoted by F indexed with the r.v.'s name. It is well-known that for model (1),

$$f_S(s) = \sum_{n=0}^{\infty} p(n) f_X^{*n}(x),$$

$$\mathbb{E}S = \mathbb{E}N\mathbb{E}X, \text{Var}S = (\mathbb{E}X)^2 \text{Var}N + \mathbb{E}N\text{Var}X,$$

where f^{*n} is the n -fold convolution of the function f , iteratively defined by

$$f^{*0}(x) = \begin{cases} 1, & x = 0 \\ 0, & \text{otherwise} \end{cases}, f^{*1} = f, f^{*(n+1)} = f * f^{*n}, \text{ with } (f * f)(x) = \int_{\mathbb{R}} f(y) f(x - y) dy.$$

In order to relax the independence condition between the number of claims and the claim amounts, we replace the above assumptions with the following ones:

Hypothesis 1. Given $N = n$, the r.v.s X_1, X_2, \dots are assumed to be i.i.d.

Hypothesis 2. Assume a Sarmanov type dependence between each X_i and N , i.e., $(X_i, N)_{i \geq 1}$ are identically distributed with the mixed Sarmanov pdf

$$f_{X,N}(x, n) = \begin{cases} p(0), & n = x = 0 \\ p(n) f(x) (1 + \omega \psi(n) \phi(x)), & n \geq 1, x > 0 \end{cases} \quad (2)$$

where ψ and ϕ are bounded non-constant kernel functions, $\omega \in \mathbb{R}$ and f is a pdf; for simplicity, we denote by Y an r.v. having pdf f and representing $X > 0$. Note that the pdf (2) is mixed because it joins the continuous pdf f and the discrete pmf p . In order for (2) to define a proper pdf, we impose the conditions

$$\sum_{n \geq 1} \psi(n) p(n) = \int_{\mathbb{R}} \phi(x) f(x) dx = 0, \text{ and} \quad (3)$$

$$1 + \omega \psi(n) \phi(x) \geq 0, \text{ for all } n \geq 1, x > 0. \quad (4)$$

With \mathcal{L}_Y denoting the Laplace transform of the r.v. Y , we shall use the exponential kernels: $\phi(y) = e^{-\gamma y} - \mathcal{L}_Y(\gamma)$ and $\psi(n) = e^{-\delta n} - \frac{\mathcal{L}_N(\delta) - p(0)}{1 - p(0)}$. Then, letting

$$\begin{aligned} m_1 &= \inf_{n \geq 1} \psi(n) = -\frac{\mathcal{L}_N(\delta) - p(0)}{1 - p(0)}, M_1 = \sup_{n \geq 1} \psi(n) = 1 - \frac{\mathcal{L}_N(\delta) - p(0)}{1 - p(0)}, \\ m_2 &= \inf_{x > 0} \phi(x) = -\mathcal{L}_Y(\gamma), M_2 = \sup_{x > 0} \phi(x) = 1 - \mathcal{L}_Y(\gamma), \end{aligned}$$

from condition (4), ω is restricted to the following interval

$$\max \left\{ -\frac{1}{m_1 m_2}, -\frac{1}{M_1 M_2} \right\} \leq \omega \leq \min \left\{ -\frac{1}{m_1 M_2}, -\frac{1}{M_1 m_2} \right\}. \quad (5)$$

Note that under the assumption H2, the distribution of the r.v. X will have both an absolutely continuous component (with pdf f_X) and a probability mass at 0; hence, the distribution of S also has a probability mass at 0 and the pdf f_S .

For the mixed Sarmanov pdf (2), it can be easily deduced that:

$$\Pr(X = 0) = p(0), f_X(x) = (1 - p(0)) f(x), x > 0; \quad (6)$$

$$\Pr(X = 0 | N = n) = \begin{cases} 1, & n = 0 \\ 0, & n \geq 1 \end{cases};$$

$$f_{X|N=n}(x) = f(x) (1 + \omega \psi(n) \phi(x)), x > 0, n \geq 1; \quad (7)$$

$$\Pr(N = n | X = x) = \begin{cases} 1, & n = x = 0 \\ \frac{p(n)}{1 - p(0)} (1 + \omega \psi(n) \phi(x)), & n \geq 1, x > 0 \end{cases}.$$

In the following proposition, we present the distribution of S . Its proof is given in the Appendix A.

Proposition 1. Under the assumptions H1-H2, it holds that

$$F_S(s) = p(0) + \sum_{n=1}^{\infty} p(n) F_{X|N=n}^{*n}(x),$$

$$\Pr(S = 0) = p(0), f_S(s) = \sum_{n=1}^{\infty} p(n) f_{X|N=n}^{*n}(x), x > 0. \quad (8)$$

To evaluate the pdf of S based on formula (8), we shall need the following result, which gives a formula for the conditional convolutions.

Proposition 2. Under the assumptions H1-H2, for $x > 0$ it holds that

$$f_{X|N=n}^{*m}(x) = f^{*m}(x) + \sum_{k=1}^m \binom{m}{k} \omega^k \psi^k(n) \left((f\phi)^{*k} * f^{*(m-k)} \right)(x).$$

Next, we present in terms of Y the first two moments of the aggregate claims S and the correlation coefficient between X and N .

Proposition 3. Under the assumptions H1-H2, the expected value and variance of S are given by

$$\begin{aligned} \mathbb{E}S &= \mathbb{E}N\mathbb{E}Y + \omega\mathbb{E}[N\psi(N)]\mathbb{E}[Y\phi(Y)], \\ \text{Var}S &= (\mathbb{E}Y)^2 \text{Var}N + \mathbb{E}N\text{Var}Y + \omega^2\mathbb{E}^2[Y\phi(Y)] \left(\text{Var}[N\psi(N)] - \mathbb{E}[N\psi^2(N)] \right) \\ &\quad + 2\omega\mathbb{E}Y\mathbb{E}[Y\phi(Y)] \left(\mathbb{E}[N^2\psi(N)] - \mathbb{E}N\mathbb{E}[N\psi(N)] - \mathbb{E}[N\psi(N)] \right) \\ &\quad + \omega\mathbb{E}[Y^2\phi(Y)]\mathbb{E}[N\psi(N)], \end{aligned}$$

while the correlation coefficient between X and N is

$$\text{corr}(X, N) = \frac{\omega\mathbb{E}[N\psi(N)]\mathbb{E}[Y\phi(Y)] + p(0)\mathbb{E}N\mathbb{E}Y}{\sqrt{(1-p(0))(\text{Var}Y + p(0)\mathbb{E}^2[Y])\text{Var}N}}.$$

Remark 1. Let $m \geq 1$ and $\mathbf{x} = (x_1, \dots, x_m)$. Then, based on H1 and (7), the conditional distribution of (X_1, X_2, \dots, X_m) given $N = n$ is

$$f_{X_1, X_2, \dots, X_m|N=n}(\mathbf{x}, n) = \prod_{j=1}^m f_{X_j|N=n}(x_j) = \prod_{j=1}^m f(x_j) (1 + \omega\psi(n)\phi(x_j)),$$

therefore, the joint distribution of $(X_1, X_2, \dots, X_m, N)$ results as

$$\begin{aligned} f_{X_1, X_2, \dots, X_m, N}(\mathbf{x}, n) &= p(n) \prod_{j=1}^m f(x_j) (1 + \omega\psi(n)\phi(x_j)) \\ &= p(n) \left(\prod_{j=1}^m f(x_j) \right) \left(1 + \sum_{k=1}^m \omega^k \psi^k(n) \sum_{j_1 < \dots < j_k} \prod_{l=1}^k \phi(x_{j_l}) \right). \end{aligned}$$

2.2. Simulation from the Collective Model

To simulate values from the distribution of S under the assumptions H1-H2, we use the inversion method for the conditional cdf of X given $N = n$, i.e., we use

$$\begin{aligned} F_{X|N=0}(0) &= 1, \\ F_{X|N=n}(x) &= F_Y(x) + \omega\psi(n) \int_0^x f(y)\phi(y)dy, n \geq 1, x > 0. \end{aligned} \tag{9}$$

Therefore, we first simulate the value n from the distribution of N . If $n = 0$ then $s = 0$; otherwise, we generate n uniform $U(0, 1)$ values $(u_i)_{i=1}^n$ and, by solving the equation $F_{X|N=n}(x_i) = u_i$ for x_i , we obtain the vector $(x_i)_{i=1}^n$. Then, $s = \sum_{i=1}^n x_i$ is a value generated from S according to model (1).

2.3. Parameters Estimation

Let $(n_i; \mathbf{x}_i)_{i=1}^K = (n_i; x_{i1}, \dots, x_{in_i})_{i=1}^K$ be a random sample of the number of claims and individual claim amounts. Let $\mathbf{0}$ be the 0 vector, let θ and ν be, respectively, the parameters vectors of the marginal distributions of N and Y , while ω is the dependence parameter of the Sarmanov distribution. From (2), the corresponding log-likelihood function is

$$\begin{aligned} \ln L \left((n_i; \mathbf{x}_i)_{i=1}^K \right) &= \sum_{\{i:n_i=0, \mathbf{x}_i=\mathbf{0}\}} \ln p(0; \theta) + \sum_{\{i:n_i \geq 1, x_{ij} > 0\}} \left[\ln p(n_i; \theta) \right. \\ &\quad \left. + \sum_{j=1}^{n_i} \ln f(x_{ij}; \nu) + \sum_{j=1}^{n_i} \ln (1 + \omega \psi(n_i) \phi(x_{ij})) \right] \\ &= \ln L \left((n_i)_{i=1}^K; \theta \right) + \ln L \left(\{ \mathbf{x}_i | x_{ij} > 0, i = 1, \dots, K, j = 1, \dots, n_i \}; \nu \right) \\ &\quad + \sum_{i=1}^K \sum_{j=1}^{n_i} \ln (1 + \omega \psi(n_i) \phi(x_{ij})), \end{aligned} \tag{10}$$

where $L \left((n_i)_{i=1}^K; \theta \right)$ is the likelihood function corresponding to the marginal r.v. N and $L \left(\{ \mathbf{x}_i | x_{ij} > 0, i = 1, \dots, K, j = 1, \dots, n_i \}; \nu \right)$ the one corresponding to Y .

Since maximizing the log-likelihood expressed in (10) is very difficult, specifically due to the close relationship that exists between the dependency parameter and the parameters associated with the marginal distributions, we define $l_{(n_i, \mathbf{x}_i)_{i=1}^K}(\theta; \nu | \omega)$ to be the log-likelihood function corresponding to the marginal parameters given ω , and $l_{(n_i, \mathbf{x}_i)_{i=1}^K}(\omega | \theta; \nu)$ the log-likelihood function of the dependence parameter given the marginal parameters θ, ν , and we determine the Maximum Likelihood Estimation (MLE) of the parameters in two phases:

Phase 1 By MLE, find initial values for the parameters of the marginal distributions. Then, iterate the following two steps:

- Step 1** (iteration j) Given the parameters for the marginal distributions, find $\hat{\omega}^j$ within the interval defined in (5) for this dependence parameter by maximizing the log-likelihood $l_{(n_i, \mathbf{x}_i)_{i=1}^K}(\omega | \theta; \nu)$;
- Step 2** Given $\hat{\omega}^j$, obtain new values for the parameters of the marginals by maximizing the log-likelihood function $l_{(n_i, \mathbf{x}_i)_{i=1}^K}(\theta; \nu | \omega)$.

Repeat steps 1 and 2 until convergence. If the dependence parameter is located at an extreme of the interval, recalculate these intervals using the parameters for marginals obtained in Step 2.

Phase 2 Starting with the initial parameters estimated in Phase 1, perform full MLE.

As a result that our parametric space is bounded, we used the `optim()` function of R with the method L-BFGS-B to perform the optimizations.

The method proposed in Phase 1 is known as the Inference From Margin (IFM) method, and has been widely used in the estimation of Copulas, see [14] for a review. Regarding the estimation of the Sarmanov distribution, the method in two phases, IFM and full MLE, has already been used with excellent results by Bolancé and Vernic [8] for the case of NB marginals.

3. Particular Cases

3.1. Particular Severity Distributions

For our particular dataset, we considered two severity distributions: Lognormal and Gamma.

Under the Gamma severity distribution assumption, $Y \sim Ga(\alpha, \beta)$, $\alpha, \beta > 0$, with β the rate parameter, we recall that

$$\mathbb{E}Y = \frac{\alpha}{\beta}, \mathbb{E}[Y^2] = \frac{\alpha(\alpha + 1)}{\beta^2}, \text{Var}Y = \frac{\alpha}{\beta^2}, \mathcal{L}_Y(\gamma) = \left(\frac{\beta}{\beta + \gamma}\right)^\alpha.$$

Furthermore, given that

$$\mathbb{E}[Ye^{-\gamma Y}] = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty y^{\alpha+1-1} e^{-(\beta+\gamma)y} dy = \frac{\alpha\beta^\alpha}{(\beta + \gamma)^{\alpha+1}},$$

and

$$\mathbb{E}[Y^2e^{-\gamma Y}] = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty y^{\alpha+2-1} e^{-(\beta+\gamma)y} dy = \frac{\alpha(\alpha + 1)\beta^\alpha}{(\beta + \gamma)^{\alpha+2}},$$

for the exponential kernel $\phi(y) = e^{-\gamma y} - \mathcal{L}_Y(\gamma)$, we easily obtain

$$\mathbb{E}[Y\phi(Y)] = -\frac{\alpha\gamma\beta^{\alpha-1}}{(\beta + \gamma)^{\alpha+1}}, \mathbb{E}[Y^2\phi(Y)] = -\frac{\alpha(\alpha + 1)\gamma\beta^{\alpha-2}(2\beta + \gamma)}{(\beta + \gamma)^{\alpha+2}}.$$

We approach the Lognormal severity distribution in a different way. We recall that if Y follows a Lognormal distribution $LN(\mu, \sigma^2)$, $\sigma > 0$, then $Z = \ln Y$ follows a normal distribution $N(\mu, \sigma^2)$. Therefore, it is easier to estimate the model having the same counting distribution, but normal severity distribution, using logarithmized claim amounts. So, we first consider the bivariate Sarmanov distribution $f_{Z,N}$ with $Z \sim N(\mu, \sigma^2)$ and exponential kernel ϕ . However, since the domain of the normal distribution is the entire real line and the kernel function ϕ must be bounded, we shall work with a left truncated normal distribution with left truncation point a , i.e., $Z \sim LTN(\mu, \sigma^2, a)$, and we select this truncation point such that $\Pr(Z \leq a)$ is almost zero; note that a good choice is $a = -3\sigma + \mu$ (another simple choice for a would be the minimum of the log-data). Hence,

$$f_{Z,N}(z, n) = p(n) f(z; \mu, \sigma^2, a) (1 + \omega\psi(n)\phi(z)), n \geq 1, z > a, \tag{11}$$

where

$$f(z; \mu, \sigma^2, a) = \frac{1}{K\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} = \frac{1}{K\sigma} \varphi\left(\frac{z-\mu}{\sigma}\right),$$

$$\phi(z) = e^{-\gamma z} - \mathcal{L}_Z(\gamma) = e^{-\gamma z} - e^{-\gamma\mu + \frac{\gamma^2\sigma^2}{2}} \frac{1 - \Phi\left(\frac{a-\mu+\gamma\sigma^2}{\sigma}\right)}{K}.$$

Here $K = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$, while φ and Φ denote the pdf and, respectively, cdf of the standard normal distribution. Then, the limits for the interval (5) of ω are

$$m_2 = \inf_{z>a} \phi(z) = -\mathcal{L}_Z(\gamma), M_2 = \sup_{z>a} \phi(z) = e^{-\gamma a} - \mathcal{L}_Z(\gamma).$$

Now, to obtain the bivariate Sarmanov distribution with a truncated Lognormal marginal, we change variable $Y = e^Z$ in (11) and have

$$f_{Y,N}(y, n) = \frac{1}{y} p(n) f(\ln y; \mu, \sigma^2, a) (1 + \omega\psi(n)\phi(\ln y))$$

$$= p(n) f_{LTLN(\mu, \sigma^2, a)}(y) \left(1 + \omega\psi(n) \left(\frac{1}{y^\gamma} - \mathcal{L}_{\ln Y}(\gamma)\right)\right), n \geq 1, y > e^a,$$

where $f_{LTLN(\mu, \sigma^2, a)}(y) = \frac{1}{K\sigma y} \varphi\left(\frac{\ln y - \mu}{\sigma}\right)$ is the pdf of the left truncated Lognormal distribution $LTLN(\mu, \sigma^2, a)$. It follows that the bivariate Sarmanov distribution with a Lognormal marginal is given by

$$f_{X,N}(x, n) = \begin{cases} p(0), & n = x = 0 \\ p(n) f_{LTLN(\mu, \sigma^2, a)}(x) (1 + \omega \psi(n) \tilde{\phi}(x)), & n \geq 1, x > e^a \end{cases}$$

where the kernel function $\tilde{\phi}(x) = x^{-\gamma} - \mathcal{L}_{LTLN(\mu, \sigma^2, a)}(\gamma)$ is not of exponential type.

The following proposition gives the needed ingredients to calculate the expected value and variance of S under the simplifying assumption $\gamma = 1$.

Proposition 4. For $Y \sim LTLN(\mu, \sigma^2; a)$ and $\gamma = 1$, it holds that

$$\begin{aligned} \mathbb{E}Y &= e^{\mu + \frac{\sigma^2}{2}} \frac{K_2}{K}, \\ \mathbb{E}[Y^2] &= e^{2\mu + 2\sigma^2} \frac{K_3}{K}, \\ \mathbb{E}[Y\tilde{\phi}(Y)] &= 1 - e^{\sigma^2} \frac{K_1 K_2}{K^2} \\ \mathbb{E}[Y^2\tilde{\phi}(Y)] &= \mathbb{E}Y - e^{\mu + \frac{5\sigma^2}{2}} \frac{K_1 K_3}{K^2} = \frac{1}{K} e^{\mu + \frac{\sigma^2}{2}} \left(K_2 - e^{2\sigma^2} \frac{K_1 K_3}{K} \right), \end{aligned}$$

where $K_1 = 1 - \Phi\left(\frac{a - \mu + \sigma^2}{\sigma}\right)$, $K_2 = 1 - \Phi\left(\frac{a - \mu - \sigma^2}{\sigma}\right)$, $K_3 = 1 - \Phi\left(\frac{a - \mu - 2\sigma^2}{\sigma}\right)$.

3.2. Particular Counting Distributions

As counting distributions, we consider the Poisson and Negative Binomial distributions. The following result holds.

Lemma 1. Let N follow a certain discrete distribution with support \mathbb{N} and let $\psi(n) = e^{-\delta n} - \hat{\mathcal{L}}_N(\delta)$ be the corresponding exponential kernel with $\hat{\mathcal{L}}_N(\delta) = \frac{\mathcal{L}_N(\delta) - p(0)}{1 - p(0)}$. Then,

$$\mathbb{E}[N\psi(N)] = \mathbb{E}[Ne^{-\delta N}] - \hat{\mathcal{L}}_N(\delta) \mathbb{E}[N], \tag{12}$$

$$\mathbb{E}[N^2\psi(N)] = \mathbb{E}[N^2e^{-\delta N}] - \hat{\mathcal{L}}_N(\delta) \mathbb{E}[N^2], \tag{13}$$

$$\mathbb{E}[N\psi^2(N)] = \mathbb{E}[Ne^{-2\delta N}] - 2\hat{\mathcal{L}}_N(\delta) \mathbb{E}[Ne^{-\delta N}] + \hat{\mathcal{L}}_N^2(\delta) \mathbb{E}[N], \tag{14}$$

$$\mathbb{E}[N^2\psi^2(N)] = \mathbb{E}[N^2e^{-2\delta N}] - 2\hat{\mathcal{L}}_N(\delta) \mathbb{E}[N^2e^{-\delta N}] + \hat{\mathcal{L}}_N^2(\delta) \mathbb{E}[N^2]. \tag{15}$$

In the following result, we present formulas needed to evaluate the expected value and variance of S given in Proposition 3, for our particular distributions.

Proposition 5. Let $\psi(n) = e^{-\delta n} - \frac{\mathcal{L}_N(\delta) - p(0)}{1 - p(0)}$ be the exponential kernel.

(i) If $N \sim Po(\lambda)$, then

$$\begin{aligned} \mathbb{E}[N\psi(N)] &= \lambda e^{-\lambda} \left(e^{\lambda e^{-\delta}-\delta} - \frac{e^{\lambda e^{-\delta}} - 1}{1 - e^{-\lambda}} \right), \\ \mathbb{E}[N^2\psi(N)] &= \lambda e^{-\lambda} \left[e^{\lambda e^{-\delta}-\delta} (\lambda e^{-\delta} + 1) - (\lambda + 1) \frac{e^{\lambda e^{-\delta}} - 1}{1 - e^{-\lambda}} \right], \\ \mathbb{E}[N\psi^2(N)] &= \lambda \left[e^{\lambda(e^{-2\delta}-1)-2\delta} + \frac{e^{\lambda e^{-\delta}} - 1}{(e^\lambda - 1)^2} \left(e^{\lambda e^{-\delta}} (1 - 2e^{-\delta} + 2e^{-\lambda-\delta}) - 1 \right) \right], \\ \mathbb{E}[N^2\psi^2(N)] &= \lambda \left[e^{\lambda(e^{-2\delta}-1)-2\delta} (\lambda e^{-2\delta} + 1) + \frac{e^{\lambda e^{-\delta}} - 1}{(e^\lambda - 1)^2} \right. \\ &\quad \left. \times \left(e^{\lambda e^{-\delta}} (\lambda + 1 - 2e^{-\delta} (\lambda e^{-\delta} + 1) (1 - e^{-\lambda})) - \lambda - 1 \right) \right]. \end{aligned}$$

(ii) If $N \sim NB(r, p)$, then

$$\begin{aligned} \mathbb{E}[N\psi(N)] &= \frac{rqp^r}{(1 - qe^{-\delta})^r} \left(\frac{1}{e^\delta - q} - \frac{1 - (1 - qe^{-\delta})^r}{p(1 - p^r)} \right), \\ \mathbb{E}[N^2\psi(N)] &= \frac{rqp^r}{(1 - qe^{-\delta})^r} \left[\frac{rq + e^\delta}{(e^\delta - q)^2} - (1 + qr) \frac{1 - (1 - qe^{-\delta})^r}{p^2(1 - p^r)} \right], \\ \mathbb{E}[N\psi^2(N)] &= rqp^r \left[\frac{e^{-2\delta}}{(1 - qe^{-2\delta})^{r+1}} + \frac{1 - (1 - qe^{-\delta})^r}{(p^{-r} - 1)(1 - qe^{-\delta})^{2r}} \left(\frac{1 - (1 - qe^{-\delta})^r}{p(1 - p^r)} - \frac{2e^{-\delta}}{1 - qe^{-\delta}} \right) \right] \\ \mathbb{E}[N^2\psi^2(N)] &= rqp^r \left[\frac{e^{-2\delta}(rqe^{-2\delta} + 1)}{(1 - qe^{-2\delta})^{r+2}} + \frac{1 - (1 - qe^{-\delta})^r}{(p^{-r} - 1)(1 - qe^{-\delta})^{2r}} \right. \\ &\quad \left. \times \left(\frac{1 - (1 - qe^{-\delta})^r}{1 - p^r} \frac{1 + qr}{p^2} - \frac{2e^{-\delta}(rqe^{-\delta} + 1)}{(1 - qe^{-\delta})^2} \right) \right]. \end{aligned}$$

In the numerical study that we present below, we combine the two discussed counting distributions with the two severity distributions and obtain the following particular compound distributions: compound Poisson–Gamma, compound Negative Binomial–Gamma, compound Poisson–Lognormal, and compound Negative Binomial–Lognormal.

4. Numerical Study

We analyzed a dataset containing a sample of $K = 99,972$ Spanish insureds with a total of 8872 claims. We assumed that they have a homogeneous risk profile. For each individual, we obtained information on the number of claims and on the individual cost of each claim notified by each insured. Our aim was to fit the bivariate Sarmanov distribution and to check the effect of dependence between frequency and severity on the risk premium. We compared the results obtained when considering that $S = N\bar{X}$, where \bar{X} represents the cost per policyholder (calculated as the mean of individual claim amounts, see comment after model (1)), and when considering that $S = \sum_{j=1}^N X_j$, where X_j represents the j -th claim amount of the policyholder or cost per claim; i.e., we compared the results obtained by taking into account the heterogeneity of the claim costs for each insured and by considering that this heterogeneity does not exist.

In Table 1, we display the results of the initial analysis of the number of claims, consisting in the basic descriptives and initially estimated parameters (by MLE) for the marginal distribution associated

with this variable. From the values of the Chi-square statistic, we can see that the best fit is obtained with the NB distribution.

In Table 2, we show the basic descriptive statistics for the cost per policyholder and for the cost per claim, together with the MLE parameters of the Gamma and Lognormal distributions for these variables. The main difference between the two variables is in the scale and shape parameters. As expected, the cost per claim r.v. has larger variance and more pronounced right skewness than the cost per policyholder. We compared the goodness of fit Akaike Information Criterion (AIC) for the Gamma and Lognormal distributions, and obtained that the best fit is provided by the Lognormal distribution.

Table 1. Number of claims: true, Poisson and Negative Binomial (NB) fitted frequencies, and Chi-square statistic. At the bottom: Maximum Likelihood Estimation (MLE) parameters included.

Number of Cases			
99,972 Policyholders			
Frequency	TRUE	Poisson	NB
0	92,538.00	91,482.28	92,524.63
1	6166.00	8118.58	6285.65
2	1122.00	360.24	950.48
3	125.00	10.66	170.11
4	18.00	0.24	32.81
5	3.00	0.00	1.73
	Chi-Square	6761.20	52.81
Initial Parameters		$\lambda = 0.0887$	$r = 0.2897$ $p = 0.7655$

Table 2. Descriptive statistics of cost variables. At the bottom: MLE parameters of Gamma and Lognormal distributions included.

	Number of Cases	Mean	Median	STD	Skewness	Pearson's Correlation
Cost per claim	8872 Claims	859.92	513.50	2448.27	24.01	0.31
Cost per policyholder	7434 Policyholders	758.13	513.50	1580.81	15.72	0.38
		Gamma		Lognormal		
Cost per claim	Initial Parameter	$\alpha = 0.6631$	$\beta = 0.0008$	$\mu = 5.8384$	$\sigma = 1.3554$	
	AIC	136,470		134,172		
Cost per policyholder	Initial Parameter	$\alpha = 0.7148$	$\beta = 0.0009$	$\mu = 5.7882$	$\sigma = 1.3441$	
	AIC	112,829		111,557		

The results of the estimated bivariate Sarmanov distributions using the procedure described in Section 2.3 are shown in Table 3. In both cases, i.e., cost per claim and cost per policyholder, the best fit was obtained with the NB–Lognormal model. The correlations ρ calculated for this type of model are consistent with the empirical correlations given in Table 2.

In Table 4, we present the values of the estimated mean and variance of S obtained when using each estimated model presented in Table 3. These values are the ones used to calculate the risk premium. If we compare the results obtained using the estimated dependence parameter $\omega > 0$ with the results obtained by assuming $\omega = 0$ (i.e., independence), we observe that, as we expected, the positive values estimated for the dependence case increase the mean and the variance. Furthermore, the largest values are obtained with the NB–Lognormal Sarmanov distribution.

To see the effect on risk premiums, in Table 5 we calculated the risk premium according to the standard deviation principle, i.e., $\pi_R = \mathbb{E}S + \delta\sqrt{VarS}$, where δ is the loading constant. We display the pure ($\delta = 0$) and risk ($\delta > 0$) premiums, assuming that $\delta = 1$; we also considered the case when N and X are independent (i.e., $\omega = 0$), and when N and X are Sarmanov distributed for both estimated

NB–Lognormal models, with cost per claim and cost per policyholder. We can see that, by using the NB–Lognormal distribution, the risk premiums are larger for the cost per policyholder model. Specifically, in the dependent case, using the individual claim cost information allows the company to reduce its risk premium with approximately 55 Euros.

Table 3. Estimation results of bivariate Sarmanov distributions using Poisson–Gamma, Poisson–Lognormal, NB–Gamma, and NB–Lognormal marginals.

Cost Per Claim							
Gamma				Lognormal			
Poisson		NB		Poisson		NB	
λ	0.0877	r	0.5998	λ	0.0875	r	1.2693
		p	0.8727			p	0.9355
α	0.6650	α	0.5892	μ	5.8384	μ	5.8384
β	0.0008	β	0.0006	σ	1.3553	σ	1.3552
ω	2.8197	ω	2.9530	ω	17.1549	ω	17.5107
ρ	0.6020	ρ	0.5625	ρ	0.3769	ρ	0.3717
AIC	199,566	AIC	199,109	AIC	197,249	AIC	195,744
BIC	199,583	BIC	199,126	BIC	197,264	BIC	195,759

Cost Per Policyholder							
Gamma				Lognormal			
Poisson		NB		Poisson		NB	
λ	0.0887	r	0.2897	λ	0.0887	r	0.2897
		p	0.7655			p	0.7655
α	0.7152	α	0.6951	μ	5.7882	μ	5.7882
β	0.0009	β	0.0009	σ	1.3441	σ	1.3441
ω	2.8157	ω	3.0631	ω	16.8899	ω	18.3588
ρ	0.6152	ρ	0.5753	ρ	0.3824	ρ	0.3621
AIC	175,690	AIC	173,754	AIC	174,402	AIC	172,374
BIC	175,707	BIC	173,769	BIC	174,417	BIC	172,389

Table 4. Expectation and variance of S obtained from the estimated Sarmanov models, and by assuming independence ($\omega = 0$).

Cost Per Claim					
Gamma			Lognormal		
	Poisson	NB	Poisson	NB	
$\omega > 0$	$E(S)$	74.64	85.96	75.40	75.46
	$V(S)$	158,977.20	240,279.00	407,210.20	412,289.50
$\omega = 0$	$E(S)$	74.62	85.89	75.33	75.33
	$V(S)$	158,880.50	239,745.70	406,504.30	411,029.60

Cost Per Policyholder					
Gamma			Lognormal		
	Poisson	NB	Poisson	NB	
$\omega > 0$	$E(S)$	67.27	68.26	71.66	71.87
	$V(S)$	128,654.30	174,034.60	378,528.60	490,725.00
$\omega = 0$	$E(S)$	67.26	68.20	71.59	71.59
	$V(S)$	128,584.80	173,651.10	377,286.30	484,877.30

Table 5. Premiums obtained with NB–Lognormal models ($\delta = 1$ for risk premiums).

	Pure Premium, $\delta = 0$		Risk Premium, $\delta = 1$	
	Indep. $\omega = 0$	Depend. $\omega > 0$	Indep. $\omega = 0$	Depend. $\omega > 0$
Cost per policyholder	71.59	71.87	767.92	772.39
Cost per claim	75.33	75.46	716.45	717.56

5. Conclusions

In this paper, we have shown how the flexibility of the Sarmanov model allows us to introduce dependence between different types of variables, discrete and continuous. We have used the Sarmanov distribution to model the bivariate distribution joining the number of claims and the individual claim amounts, which is needed to estimate, e.g., the moments of the aggregate claims in the collective risk model considering dependence between its variables. We numerically compared the results obtained using the cost per claim and also using the cost per policyholder; in both cases, the NB–Lognormal Sarmanov model proved to provide the best fit to our real dataset.

We have also analyzed the differences between the expectations and the variances of the aggregate claims in the collective models obtained using alternative estimated distributions. We note that, as expected, these values are larger for the models in which a Lognormal marginal distribution is considered for the cost.

Author Contributions: Both authors contributed equally to this work in terms of conceptualization, methodology, validation, writing; software, C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We thank Fundación BBVA, Equipos de Investigación Científica en Big Data 2018 for support. Also, we are very grateful to the three referees for their nice reviews and valuable comments that helped to improve the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- MDPI Multidisciplinary Digital Publishing Institute
- DOAJ Directory of open access journals
- TLA Three letter acronym
- LD linear dichroism

Appendix A

Proof of Proposition 1. Since $S = 0$ when $N = 0$, we clearly obtain that $\Pr(S = 0) = p(0)$. Then, for $x > 0$, we have

$$F_S(x) = \Pr(S \leq s) = \Pr(S = 0) + \sum_{n=1}^{\infty} \Pr\left(\sum_{j=1}^n X_j \leq x \mid N = n\right) p(n) = p(0) + \sum_{n=1}^{\infty} p(n) F_{X|N=n}^*(x).$$

Differentiating with respect to x easily yields the formula of f_S . \square

Proof of Proposition 2. Let $x > 0$. We prove the result by induction: when $m = 1$, based on the definition of f^{*0} , we have

$$f_{X|N=n}^{*1}(x) = f_{X|N=n}(x) = f(x) + \omega\psi(n)(f\phi)(x) = f(x) + \omega\psi(n)\left((f\phi) * f^{*0}\right)(x).$$

Assuming that the result holds for $m \geq 1$ and taking $m + 1$, we obtain

$$\begin{aligned}
 f_{X|N=n}^{*(m+1)}(x) &= \int_D f_{X|N=n}^{*m}(y) f_{X|N=n}(x-y) dy \\
 &= \int_D \left[f^{*m}(y) + \sum_{k=1}^m \binom{m}{k} \omega^k \psi^k(n) \left((f\phi)^{*k} * f^{*(m-k)} \right) (y) \right] f(x-y) (1 + \omega\psi(n)\phi(x-y)) dy \\
 &= \int_D f^{*m}(y) f(x-y) (1 + \omega\psi(n)\phi(x-y)) dy \\
 &\quad + \sum_{k=1}^m \binom{m}{k} \omega^k \psi^k(n) \int_D \left((f\phi)^{*k} * f^{*(m-k)} \right) (y) f(x-y) (1 + \omega\psi(n)\phi(x-y)) dy \\
 &= f^{*(m+1)}(x) + \omega\psi(n) \left((f\phi) * f^{*m} \right) (x) + \sum_{k=1}^m \binom{m}{k} \omega^k \psi^k(n) \left((f\phi)^{*k} * f^{*(m+1-k)} \right) (x) \\
 &\quad + \sum_{k=1}^m \binom{m}{k} \omega^{k+1} \psi^{k+1}(n) \left((f\phi)^{*(k+1)} * f^{*(m-k)} \right) (x),
 \end{aligned}$$

where D is the integration domain. Changing the index in the last sum, we have

$$\sum_{k=1}^m \binom{m}{k} \omega^{k+1} \psi^{k+1}(n) \left((f\phi)^{*(k+1)} * f^{*(m-k)} \right) (x) = \sum_{k=2}^{m+1} \binom{m}{k-1} \omega^k \psi^k(n) \left((f\phi)^{*k} * f^{*(m+1-k)} \right) (x).$$

Inserting this into the above formula of $f_{X|N=n}^{*(m+1)}$ and rearranging gives

$$\begin{aligned}
 f_{X|N=n}^{*(m+1)}(x) &= f^{*(m+1)}(x) + \sum_{k=1}^m \left[\binom{m}{k} + \binom{m}{k-1} \right] \omega^k \psi^k(n) \left((f\phi)^{*k} * f^{*(m+1-k)} \right) (x) \\
 &\quad + \omega^{m+1} \psi^{m+1}(n) \left((f\phi)^{*(m+1)} * f^{*0} \right) (x) \\
 &= f^{*(m+1)}(x) + \sum_{k=1}^m \binom{m+1}{k} \omega^k \psi^k(n) \left((f\phi)^{*k} * f^{*(m+1-k)} \right) (x) \\
 &\quad + \omega^{m+1} \psi^{m+1}(n) \left((f\phi)^{*(m+1)} * f^{*0} \right) (x),
 \end{aligned}$$

which immediately yields the result. This completes the proof. \square

Proof of Proposition 3. We prove the expected value and variance formulas in the usual way. Using (7), we have

$$\mathbb{E}S = \mathbb{E} \left[\mathbb{E} \left[\sum_{j=1}^N X_j \mid N \right] \right] = \mathbb{E} [N\mathbb{E} [X|N]] = \mathbb{E} [N(\mathbb{E}Y + \omega\psi(N)\mathbb{E} [Y\phi(Y)])],$$

which easily yields the formula of $\mathbb{E}S$. In what concerns the variance, we use

$$VarS = Var [\mathbb{E} [S|N]] + \mathbb{E} [Var [S|N]]. \tag{A1}$$

We shall need the following formulas. From (7), we have

$$\begin{aligned}
 Var [X|N] &= \mathbb{E} [X^2|N] - \mathbb{E}^2 [X|N] \\
 &= \mathbb{E} [Y^2] + \omega\psi(N)\mathbb{E} [Y^2\phi(Y)] - (\mathbb{E}Y + \omega\psi(N)\mathbb{E} [Y\phi(Y)])^2,
 \end{aligned}$$

formula that we insert into

$$\begin{aligned} \mathbb{E} [Var [S|N]] &= \mathbb{E} \left[Var \left[\sum_{j=1}^N X_j \middle| N \right] \right] = \mathbb{E} [NVar [X|N]] \\ &= \mathbb{E} [N] VarY + \omega \mathbb{E} [N\psi(N)] \mathbb{E} [Y^2\phi(Y)] - 2\omega \mathbb{E} [N\psi(N)] \mathbb{E} [Y\phi(Y)] \mathbb{E} Y \\ &\quad - \omega^2 \mathbb{E} [N\psi^2(N)] \mathbb{E}^2 [Y\phi(Y)]. \end{aligned}$$

On the other hand,

$$\begin{aligned} Var [\mathbb{E} [S|N]] &= Var [N (\mathbb{E} Y + \omega\psi(N) \mathbb{E} [Y\phi(Y)])] \\ &= \mathbb{E} [N^2 (\mathbb{E} Y + \omega\psi(N) \mathbb{E} [Y\phi(Y)])^2] - \mathbb{E}^2 [N (\mathbb{E} Y + \omega\psi(N) \mathbb{E} [Y\phi(Y)])] \\ &= \mathbb{E} [N^2] \mathbb{E}^2 [Y] + 2\omega \mathbb{E} [N^2\psi(N)] \mathbb{E} [Y\phi(Y)] \mathbb{E} Y + \omega^2 \mathbb{E} [N^2\psi^2(N)] \mathbb{E}^2 [Y\phi(Y)] \\ &\quad - \mathbb{E}^2 [N] \mathbb{E}^2 [Y] - 2\omega \mathbb{E} [N\psi(N)] \mathbb{E} [Y\phi(Y)] \mathbb{E} N \mathbb{E} Y - \omega^2 \mathbb{E}^2 [N\psi(N)] \mathbb{E}^2 [Y\phi(Y)] \\ &= \mathbb{E}^2 [Y] VarN + 2\omega \mathbb{E} Y \mathbb{E} [Y\phi(Y)] \left(\mathbb{E} [N^2\psi(N)] - \mathbb{E} N \mathbb{E} [N\psi(N)] \right) \\ &\quad + \omega^2 \mathbb{E}^2 [Y\phi(Y)] \left(\mathbb{E} [N^2\psi^2(N)] - \mathbb{E}^2 [N\psi(N)] \right), \end{aligned}$$

and after inserting all these into (A1), we obtain the variance formula.

From formula (6) it is easy to check that

$$\begin{aligned} \mathbb{E} X &= (1 - p(0)) \mathbb{E} Y, \quad \mathbb{E} [X^2] = (1 - p(0)) \mathbb{E} [Y^2], \\ VarX &= (1 - p(0)) \left(VarY + p(0) (\mathbb{E} Y)^2 \right). \end{aligned}$$

On the other hand, we have that

$$\begin{aligned} \mathbb{E} [XN] &= \sum_{n=1}^{\infty} np(n) \int_0^{\infty} xf(x) (1 + \omega\psi(n)\phi(x)) dx \\ &= \sum_{n=1}^{\infty} np(n) \int_0^{\infty} xf(x) dx + \omega \sum_{n=1}^{\infty} np(n)\psi(n) \int_0^{\infty} xf(x)\phi(x) dx \\ &= \mathbb{E} N \mathbb{E} Y + \omega \mathbb{E} [N\psi(N)] \mathbb{E} [Y\phi(Y)], \end{aligned}$$

hence

$$\begin{aligned} cov(X, N) &= \mathbb{E} [XN] - \mathbb{E} X \mathbb{E} N = \mathbb{E} N \mathbb{E} Y + \omega \mathbb{E} [N\psi(N)] \mathbb{E} [Y\phi(Y)] - (1 - p(0)) \mathbb{E} Y \mathbb{E} N \\ &= \omega \mathbb{E} [N\psi(N)] \mathbb{E} [Y\phi(Y)] + p(0) \mathbb{E} Y \mathbb{E} N, \end{aligned}$$

which, together with the above formula of $VarX$, immediately yields the stated formula of $corr(X, N)$. This completes the proof. □

Proof of Proposition 4. We note that

$$\begin{aligned} \mathbb{E} Y &= \mathbb{E} [e^Z] = \mathcal{L}_Z(-1) = e^{\mu + \frac{\sigma^2}{2}} \frac{1 - \Phi\left(\frac{a - \mu - \sigma^2}{\sigma}\right)}{K}, \\ \mathbb{E} [Y^2] &= \mathbb{E} [e^{2Z}] = \mathcal{L}_Z(-2) = e^{2\mu + \frac{4\sigma^2}{2}} \frac{1 - \Phi\left(\frac{a - \mu - 2\sigma^2}{\sigma}\right)}{K}, \end{aligned}$$

yielding the first two formulas. The other two formulas result from

$$\begin{aligned} \mathbb{E}[Y\tilde{\phi}(Y)] &= \mathbb{E}\left[Y\left(Y^{-1} - \mathcal{L}_Z(1)\right)\right] = 1 - e^{-\mu + \frac{\sigma^2}{2}} \frac{1 - \Phi\left(\frac{a - \mu + \sigma^2}{\sigma}\right)}{K} \mathbb{E}Y \\ &= 1 - e^{-\mu + \frac{\sigma^2}{2}} e^{\mu + \frac{\sigma^2}{2}} \frac{K_1}{K} \frac{K_2}{K}, \\ \mathbb{E}[Y^2\tilde{\phi}(Y)] &= \mathbb{E}\left[Y^2\left(Y^{-1} - \mathcal{L}_Z(1)\right)\right] = \mathbb{E}Y - e^{-\mu + \frac{\sigma^2}{2}} \frac{K_1}{K} \mathbb{E}[Y^2] \\ &= e^{\mu + \frac{\sigma^2}{2}} \frac{K_2}{K} - e^{-\mu + \frac{\sigma^2}{2}} \frac{K_1}{K} e^{2\mu + 2\sigma^2} \frac{K_3}{K}, \end{aligned}$$

and the result is immediate. \square

Proof of Lemma 1. The formulas (12) and (13) can be obtained directly. Formula (14) easily results from

$$\mathbb{E}\left[N\psi^2(N)\right] = \mathbb{E}\left[N\left(e^{-2\delta N} - 2\hat{\mathcal{L}}_N(\delta)e^{-\delta N} + \hat{\mathcal{L}}_N^2(\delta)\right)\right],$$

which is also the case with formula (15). \square

Proof of Proposition 5. (i) Let $N \sim Po(\lambda)$. Tamraz and Vernic [15] proved in Lemma 4.1 that $\mathbb{E}[Ne^{-\delta N}] = \lambda e^{\lambda(e^{-\delta}-1)-\delta}$, hence, applying also formula (12),

$$\mathbb{E}[N\psi(N)] = \lambda e^{\lambda(e^{-\delta}-1)-\delta} - \lambda \frac{e^{\lambda(e^{-\delta}-1)} - e^{-\lambda}}{1 - e^{-\lambda}} = \lambda e^{-\lambda} \left(e^{\lambda e^{-\delta} - \delta} - \frac{e^{\lambda e^{-\delta}} - 1}{1 - e^{-\lambda}} \right).$$

For the second formula, we use

$$\begin{aligned} \mathbb{E}\left[N^2e^{-\delta N}\right] &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{n^2\lambda^n}{n!} e^{-\delta n} = e^{-\lambda} \sum_{n=1}^{\infty} \frac{(n-1+1)(\lambda e^{-\delta})^n}{(n-1)!} \\ &= e^{-\lambda} \left[(\lambda e^{-\delta})^2 \sum_{n=2}^{\infty} \frac{(\lambda e^{-\delta})^{n-2}}{(n-2)!} + \lambda e^{-\delta} \sum_{n=1}^{\infty} \frac{(\lambda e^{-\delta})^{n-1}}{(n-1)!} \right] \\ &= e^{-\lambda} \left((\lambda e^{-\delta})^2 e^{\lambda e^{-\delta}} + \lambda e^{-\delta} e^{\lambda e^{-\delta}} \right) = \lambda e^{\lambda e^{-\delta} - \lambda - \delta} (\lambda e^{-\delta} + 1), \end{aligned}$$

that we insert into (13) and obtain

$$\begin{aligned} \mathbb{E}\left[N^2\psi(N)\right] &= \lambda e^{\lambda e^{-\delta} - \lambda - \delta} (\lambda e^{-\delta} + 1) - \lambda(\lambda + 1) \frac{e^{\lambda(e^{-\delta}-1)} - e^{-\lambda}}{1 - e^{-\lambda}} \\ &= \lambda e^{-\lambda} \left[e^{\lambda e^{-\delta} - \delta} (\lambda e^{-\delta} + 1) - (\lambda + 1) \frac{e^{\lambda e^{-\delta}} - 1}{1 - e^{-\lambda}} \right]. \end{aligned}$$

Note that $\mathbb{E}[Ne^{-2\delta N}] = \lambda e^{\lambda(e^{-2\delta}-1)-2\delta}$. Inserting into (14), we have

$$\begin{aligned} \mathbb{E}\left[N\psi^2(N)\right] &= \lambda e^{\lambda(e^{-2\delta}-1)-2\delta} + \frac{e^{\lambda(e^{-\delta}-1)} - e^{-\lambda}}{1 - e^{-\lambda}} \left(\frac{e^{\lambda(e^{-\delta}-1)} - e^{-\lambda}}{1 - e^{-\lambda}} \lambda - 2\lambda e^{\lambda(e^{-\delta}-1)-\delta} \right) \\ &= \lambda \left[e^{\lambda(e^{-2\delta}-1)-2\delta} + \frac{e^{\lambda e^{-\delta}} - 1}{(e^{\lambda} - 1)^2} \left(e^{\lambda e^{-\delta}} - 1 - 2e^{\lambda(e^{-\delta}-1)-\delta} (e^{\lambda} - 1) \right) \right] \\ &= \lambda \left[e^{\lambda(e^{-2\delta}-1)-2\delta} + \frac{e^{\lambda e^{-\delta}} - 1}{(e^{\lambda} - 1)^2} \left(e^{\lambda e^{-\delta}} - 1 - 2e^{\lambda e^{-\delta} - \delta} + 2e^{\lambda(e^{-\delta}-1)-\delta} \right) \right], \end{aligned}$$

which easily gives the third formula of (i). For the fourth formula of this case, using $\mathbb{E} [N^2 e^{-\delta N}] = \lambda e^{\lambda e^{-\delta} - \lambda - \delta} (\lambda e^{-\delta} + 1)$ and (15), we write

$$\begin{aligned} \mathbb{E} [N^2 \psi^2(N)] &= \lambda e^{\lambda(e^{-2\delta}-1)-2\delta} (\lambda e^{-2\delta} + 1) + \frac{e^{\lambda(e^{-\delta}-1)} - e^{-\lambda}}{1 - e^{-\lambda}} \\ &\quad \times \left(\frac{e^{\lambda(e^{-\delta}-1)} - e^{-\lambda}}{1 - e^{-\lambda}} \lambda (\lambda + 1) - 2\lambda e^{\lambda(e^{-\delta}-1)-\delta} (\lambda e^{-\delta} + 1) \right) \\ &= \lambda \left[e^{\lambda(e^{-2\delta}-1)-2\delta} (\lambda e^{-2\delta} + 1) + \frac{e^{\lambda e^{-\delta}} - 1}{(e^\lambda - 1)^2} \right. \\ &\quad \left. \times \left((e^{\lambda e^{-\delta}} - 1) (\lambda + 1) - 2e^{\lambda(e^{-\delta}-1)-\delta} (\lambda e^{-\delta} + 1) (e^\lambda - 1) \right) \right], \end{aligned}$$

from where the stated result.

(ii) Let $N \sim NB(r, p)$. From Tamraz and Vernic [15] we have that $\mathbb{E} [N e^{-\delta N}] = \frac{rqp^r e^{-\delta}}{(1 - qe^{-\delta})^{r+1}}$, so that using (12), we obtain

$$\begin{aligned} \mathbb{E} [N\psi(N)] &= \frac{rqp^r e^{-\delta}}{(1 - qe^{-\delta})^{r+1}} - \frac{rq \left(\frac{p}{1 - qe^{-\delta}}\right)^r - p^r}{p(1 - p^r)} \\ &= \frac{rqp^r}{(1 - qe^{-\delta})^r} \left(\frac{e^{-\delta}}{1 - qe^{-\delta}} - \frac{1 - (1 - qe^{-\delta})^r}{p(1 - p^r)} \right), \end{aligned}$$

yielding the first formula. To obtain the second stated formula, we first evaluate

$$\begin{aligned} \mathbb{E} [N^2 e^{-\delta N}] &= \sum_{n=0}^{\infty} \frac{\Gamma(r+n)}{n! \Gamma(r)} n^2 p^r (qe^{-\delta})^n = \sum_{n=1}^{\infty} \frac{\Gamma(r+n)(n-1+1)}{(n-1)! \Gamma(r)} p^r (qe^{-\delta})^n \\ &= \frac{p^r}{(1 - qe^{-\delta})^r} \left[\sum_{n=2}^{\infty} \frac{\Gamma(r+n)}{(n-2)! \Gamma(r)} (1 - qe^{-\delta})^r (qe^{-\delta})^n \right. \\ &\quad \left. + \sum_{n=1}^{\infty} \frac{\Gamma(r+n)}{(n-1)! \Gamma(r)} (1 - qe^{-\delta})^r (qe^{-\delta})^n \right] \\ &= \frac{p^r}{(1 - qe^{-\delta})^r} \left[\frac{r(r+1)(qe^{-\delta})^2}{(1 - qe^{-\delta})^2} + \frac{rqp^r e^{-\delta}}{1 - qe^{-\delta}} \right] \\ &= \frac{rqp^r e^{-\delta} (rqp^r e^{-\delta} + 1)}{(1 - qe^{-\delta})^{r+2}}. \end{aligned}$$

Therefore, based on (13), we have

$$\begin{aligned} \mathbb{E} [N^2 \psi(N)] &= \frac{rqp^r e^{-\delta} (rqp^r e^{-\delta} + 1)}{(1 - qe^{-\delta})^{r+2}} - \frac{rq(1 + qr) \left(\frac{p}{1 - qe^{-\delta}}\right)^r - p^r}{p^2(1 - p^r)} \\ &= \frac{rqp^r}{(1 - qe^{-\delta})^r} \left[\frac{e^{-\delta} (rqp^r e^{-\delta} + 1)}{(1 - qe^{-\delta})^2} - \frac{1 + qr}{p^2} \frac{1 - (1 - qe^{-\delta})^r}{1 - p^r} \right], \end{aligned}$$

which easily yields the stated formula.

Now, using (14), we obtain

$$\begin{aligned} \mathbb{E} \left[N\psi^2(N) \right] &= \frac{rqp^r e^{-2\delta}}{(1 - qe^{-2\delta})^{r+1}} + \frac{\left(\frac{p}{1 - qe^{-\delta}}\right)^r - p^r}{1 - p^r} \left(\frac{rq \left(\frac{p}{1 - qe^{-\delta}}\right)^r - p^r}{p(1 - p^r)} - \frac{2rqp^r e^{-\delta}}{(1 - qe^{-\delta})^{r+1}} \right) \\ &= rqp^r \left[\frac{e^{-2\delta}}{(1 - qe^{-2\delta})^{r+1}} + p^r \frac{1 - (1 - qe^{-\delta})^r}{(1 - p^r)(1 - qe^{-\delta})^r} \right. \\ &\quad \left. \times \left(\frac{1 - (1 - qe^{-\delta})^r}{p(1 - p^r)(1 - qe^{-\delta})^r} - \frac{2e^{-\delta}}{(1 - qe^{-\delta})^{r+1}} \right) \right], \end{aligned}$$

yielding the third formula of (ii). Inserting now $\mathbb{E} [N^2 e^{-\delta N}] = \frac{rqp^r e^{-\delta} (rqe^{-\delta} + 1)}{(1 - qe^{-\delta})^{r+2}}$ into (15) gives

$$\begin{aligned} \mathbb{E} \left[N^2 \psi^2(N) \right] &= \frac{rqp^r e^{-2\delta} (rqe^{-2\delta} + 1)}{(1 - qe^{-2\delta})^{r+2}} + \frac{p^r (1 - (1 - qe^{-\delta})^r)}{(1 - p^r)(1 - qe^{-\delta})^r} \\ &\quad \times \left(\frac{p^r (1 - (1 - qe^{-\delta})^r)}{(1 - p^r)(1 - qe^{-\delta})^r} \frac{rq(1 + qr)}{p^2} - 2 \frac{rqp^r e^{-\delta} (rqe^{-\delta} + 1)}{(1 - qe^{-\delta})^{r+2}} \right), \end{aligned}$$

and the result is immediate. This completes the proof. \square

References

1. Erhardt, V.; Czado, C. Modeling dependent yearly claim totals including zero claims in private health insurance. *Scand. Actuar. J.* **2012**, *2*, 106–129.
2. Czado, C.; Kastenmeier, R.; Brechmann, E.C.; Min, A. A mixed copula model for insurance claims and claim sizes. *Scand. Actuar. J.* **2012**, *4*, 278–305. [\[CrossRef\]](#)
3. Krämer, N.; Brechmann, E.; Silvestrini, D.; Czado, C. Total loss estimation using copula-based regression models. *Insur. Math. Econ.* **2013**, *53*, 829–839. [\[CrossRef\]](#)
4. Lee, G.Y.; Shi, P. A dependent frequency–severity approach to modeling longitudinal insurance claims. *Insur. Math. Econ.* **2019**, *87*, 115–129. [\[CrossRef\]](#)
5. Oh, R.; Shi, P.; Ahn, J.Y. Bonus-Malus premiums under the dependent frequency-severity modeling. *Scand. Actuar. J.* **2020**, *3*, 172–195. [\[CrossRef\]](#)
6. Bahraoui, Z.; Bolancé, C.; Pelican, E.; Vernic, R. On the bivariate distribution and copula. An application on insurance data using truncated marginal distributions. *Stat. Oper. Res. Trans.* **2015**, *39*, 209–230.
7. Abdallah, A.; Boucher, J.; Cossette, H. Sarmanov family of multivariate distributions for bivariate dynamic claim counts model. *Insur. Math. Econ.* **2016**, *68*, 120–133. [\[CrossRef\]](#)
8. Bolancé, C.; Vernic, R. Multivariate count data generalized linear models: Three approaches based on the Sarmanov distribution. *Insur. Math. Econ.* **2019**, *85*, 89–103. [\[CrossRef\]](#)
9. Yang, Y.; Yuen, K.C. Finite-time and infinite-time ruin probabilities in a two-dimensional delayed renewal risk model with Sarmanov dependent claims. *J. Math. Anal. Appl.* **2016**, *442*, 600–625. [\[CrossRef\]](#)
10. Guo, F.; Wang, D.; Yang, H. Asymptotic results for ruin probability in a two-dimensional risk model with stochastic investment returns. *J. Comput. Appl. Math.* **2017**, *325*, 198–221. [\[CrossRef\]](#)
11. Ratovomirija, G. On mixed Erlang reinsurance risk: Aggregation, capital allocation and default risk. *Eur. Actuar. J.* **2016**, *6*, 149–175. [\[CrossRef\]](#)
12. Ratovomirija, G.; Tamraz, M.; Vernic, R. On some multivariate Sarmanov mixed Erlang reinsurance risks: Aggregation and capital allocation. *Insur. Math. Econ.* **2017**, *74*, 197–209. [\[CrossRef\]](#)
13. Vernic, R. Capital allocation for Sarmanov’s class of distributions. *Methodol. Comput. Appl. Probab.* **2017**, *19*, 311–330. [\[CrossRef\]](#)

14. Joe, H.; Xu, J.J. The Estimation Method of Inference Functions for Margins for Multivariate Models. 1996. Available online: <https://open.library.ubc.ca/cIRcle/collections/facultyresearchandpublications/52383/items/1.0225985> (accessed on 1 August 2020).
15. Tamraz, M.; Vernic, R. On the evaluation of multivariate compound distributions with continuous severity distributions and Sarmanov's counting distribution. *ASTIN Bull.* **2018**, *48*, 841–870. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).