ON A MULTIVARIATE AGGREGATE CLAIMS MODEL WITH MULTIVARIATE POISSON COUNTING DISTRIBUTION

Elena-Grătiela ROBE-VOINEA1, Raluca VERNIC2

1 University of Bucharest, Faculty of Mathematics and Computer Science
2 “Ovidius” University of Constanţa, Faculty of Mathematics and Computer Science, and Institute for Mathematical Statistics and Applied Mathematics, Bucharest

Corresponding author: Elena-Grătiela ROBE-VOINEA, E-mail: elena_robe@yahoo.com

Abstract. In this paper, we consider a multivariate aggregate claims model that includes different types of claims from which some could simultaneously affect an insurance portfolio. For this model, we present a recursive formula to evaluate its probability function when the number of claims follows a multivariate Poisson distribution. A possible extension is also suggested.

Key words: insurance model, multivariate aggregate claims, multivariate Poisson distribution, recursion.

1. INTRODUCTION

Modelling aggregate claims is a very important task for an insurance company due to their applications in premiums calculations, reserves evaluation, reinsurance covers study etc. Therefore, there is an increasing amount of literature related to this topic, see, e.g., the book [1] and the references therein, or, more recently, the paper [2], which proposes a new distribution family for modelling insurance data (based on generalizing the exponential-Poisson distribution), or the work [3], in which the problem of existence of the restricted optimal retention in a stop-loss reinsurance is studied and the results are applied in the case of generalized Pareto distributed aggregate claims.

The main models for aggregate claims are the individual and collective ones. In this paper, we deal with the collective model in a multivariate setting motivated by the fact that some stochastic sources (like, e.g., fires, floods, traffic accidents, earthquakes) may cause different types of dependent claims. Therefore, we consider the following multivariate aggregate claims model

\[ (S_1,\ldots,S_m) = \left( \sum_{l=0}^{N_1} U_{1l} + \sum_{k=0}^{N_k} L_{1k}, \ldots, \sum_{l=0}^{N_1} U_{ml} + \sum_{k=0}^{N_m} L_{mk} \right), \] (1)

where \( m \geq 2 \) is the number of different types of claims affecting the portfolio, \( S_k \) denotes the aggregate claims of type \( k \), \( N_k \) the number of claims of only type \( k \), \( N_0 \) the number of common claims (e.g., accidents that causes all \( m \) types of different claims). Each set of claim sizes \( (U_{jl})_{j,l} \) are positive, independent and identically distributed (i.i.d.) as the generic random variable (r.v.) \( U_j \), \( 1 \leq j \leq m \), independent of the claim numbers and of the other claim sizes, including \( (L_{lk})_{k,l} \). The random vectors \( (L_{lk})_{k,l} \) are non-negative i.i.d. as the generic \( (L_1,\ldots,L_m) \), and independent of the claim numbers. Clearly, \( U_{j0} = L_{j0} = 0, \forall j \).

In the following, by a bold faced letter we denote a vector, i.e., \( X = (X_1,\ldots,X_m) \) or \( x = (x_1,\ldots,x_m) \). We shall work with discrete r.v.’s (if the claim sizes distributions are continuous, they should be discretized using, e.g., the rounding method, see [1]). If \( f \) is a probability function (p.f.), we denote by \( f^{*n} \) its \( n \)-fold convolution corresponding to the distribution of the sum of \( n \) i.i.d. r.v.’s having p.f. \( f \); note that \( f^{*1} = f \) and, by convention, \( f^{*0}(x) = \begin{cases} 1, x = 0 \\ 0, x \neq 0 \end{cases} \). Let \( f_S \) denote the p.f. of \( S \), \( f_j \) the p.f. of \( U_j \), \( 1 \leq j \leq m \), \( f_0 \) the p.f. of \( L \), and \( p \) the p.f. of \( N = (N_0,\ldots,N_m) \). Then, from (1), we easily obtain
where \( \mathbf{0} = (0, \ldots, 0) \), while the inequality \( \mathbf{x} \geq \mathbf{0} \) and the difference \( \mathbf{x} - \mathbf{k} \) are componentwise. Note that due to the convolutions in (2), \( f_S \) could be difficult to evaluate and time-consuming; therefore, alternative methods have been developed, from which recursions play an important role in actuarial mathematics (for details on alternative methods see [1], while for details on recursions see [4]).

We recall the fact that the distribution of \( N \) is also called counting distribution, while the distribution of \( S \) is called compound.

In this paper, in next section, we shall present a recursion to evaluate \( f_S \) for model (1) when the number of claims \( N \) follows a multivariate Poisson distribution. In the bivariate setting, when \( m = 2 \), a recursion for this case has been recently presented in [5]; our recursion extends this one to a general \( m \). In the simpler univariate case (\( m = 1 \)), the history of recursions involving Poisson counting distributions starts in insurance with [6] and [7], and continues with more complex recursions for compound mixed Poisson distributions discussed in [8, 9] among others, or, in the multivariate case, in [10], etc.

## 2. MAIN RESULT

### 2.1. Probability generating function

Let us denote the p.f. of a discrete random vector \( \mathbf{X} \) by \( f_X \) and its probability generating function (p.g.f.) by \( g_X \); we recall that

\[
g_X(t) = E \left[ \prod_{j=1}^{m} t_j^{X_j} \right].
\]

Moreover, as a property of the p.g.f., it holds that

\[
g_X(t) = \sum_{\mathbf{x} \geq \mathbf{0}} f_X(\mathbf{x}) \prod_{j=1}^{m} t_j^{x_j},
\]

and clearly, \( g_X(\mathbf{0}) = f_X(\mathbf{0}) \). In particular, if \( N = (N_1, \ldots, N_m) \) follows the multivariate Poisson distribution \( MPo_{m+1}(\lambda, \lambda_0, \ldots, \lambda_m) \) with \( \lambda > 0, \lambda_j > 0, \forall j \), then, from [11] we have that

\[
g_N(t) = \exp \left\{ \lambda \left( \prod_{j=0}^{m} t_j - 1 \right) + \sum_{j=0}^{m} \lambda_j (t_j - 1) \right\}.
\]

The following result holds.

**PROPOSITION 2.1.** *Under the assumptions of model (1), the p.g.f. of \( S \) is given by*

\[
g_S(t) = g_N \left( g_{L}(t), g_{C_1}(t), \ldots, g_{C_m}(t) \right).
\]

**Proof.** We have that

\[
g_S(t) = E \left[ \prod_{j=1}^{m} t_j^{S_j} \right] = E \left[ \prod_{j=1}^{m} t_j \left( \sum_{i=0}^{N_j \times U_j} \sum_{k=0}^{N_j \times L_j} \right) \left( N_0, \ldots, N_m \right) \right] \]

\[
= E \left[ \prod_{j=1}^{m} E \left[ t_j^{U_j} \right]^{N_j} \right] E \left[ \prod_{j=1}^{m} t_j^{L_j} \right]^{N_a} \right] = \left[ g_L^{N_a}(t) \prod_{j=1}^{m} g_{U_j}(t_j) \right],
\]

which immediately yields formula (5) and completes the proof.
2.2. The recursion

We shall now obtain a recursion for the p.f. \( f_s \) for model (1) when \( N \) follows a multivariate Poisson distribution.

**PROPOSITION 2.2.** Under the assumptions of model (1), if \( N \) follows the multivariate Poisson distribution \( \text{MPo}_{m+1}(\lambda, \lambda_0, \ldots, \lambda_m) \), then the p.f. of \( S \) satisfies the recursion

\[
f_s(x) = \frac{\lambda_k}{x_k} \sum_{z_k=0}^{x_k} f_k(z_k) f_s(x_1, \ldots, x_{k-1}, x_k - z_k, x_{k+1}, \ldots, x_m) + \frac{\lambda_m}{x_m} \sum_{z_m=0}^{x_m} z_m f_k(z) f_s(x - z) + \frac{\lambda_k}{x_k} \sum_{v_k=0}^{x_k} \prod_{i=1}^{m} f_i(u_i) f_0(v - u) f_s(x - v), \quad x_k \geq 1, \ 1 \leq k \leq m,
\]

with starting value

\[
f_s(0) = \exp \left\{ \lambda \left( f_0(0) \prod_{j=1}^{m} f_j(0) - 1 \right) + \sum_{j=1}^{m} \lambda_j \left( f_j(0) - 1 \right) + \lambda_0 (f_0(0) - 1) \right\}.
\]

**Proof.** We start by inserting the p.g.f. (4) of \( N \) into (5) and obtain

\[
g_s(t) = \exp \left\{ \lambda \left( g_L(t) \prod_{j=1}^{m} g_{U_j}(t_j) - 1 \right) + \sum_{j=1}^{m} \lambda_j \left( g_{U_j}(t_j) - 1 \right) + \lambda_0 (g_L(t) - 1) \right\}.
\]

Taking here \( t = 0 \) and using the properties of the p.g.f., we easily get the starting value (7).

To obtain the recursion, we recall from (3) that

\[
g_s(t) = \sum_{x \geq 0} f_s(x) \prod_{i=1}^{m} t_i^{x_i}, \quad g_L(t) = \sum_{x \geq 0} f_0(x) \prod_{i=1}^{m} t_i^{x_i}, \quad g_{U_j}(t_j) = \sum_{x \geq 0} f_k(x) t_j^{x_j},
\]

which yields

\[
\frac{\partial g_s(t)}{\partial t_k} = \sum_{x \geq 0} f_s(x) x_k t_k^{x_k - 1} \prod_{i=1, i \neq k}^{m} t_i^{x_i}, \quad \frac{\partial g_s(t)}{\partial t_k} = \sum_{x \geq 0} f_0(x) x_k t_k^{x_k - 1} \prod_{i=1, i \neq k}^{m} t_i^{x_i}, \quad g_{U_j}(t_j) = \sum_{x \geq 0} f_k(x) t_j^{x_j}.
\]

We now consider the partial derivative of \( g_s \) with respect to \( t_k, 1 \leq k \leq m \) i.e.,

\[
\frac{\partial g_s(t)}{\partial t_k} = g_s(t) \left[ \lambda \left( \frac{\partial g_L(t)}{\partial t_k} \prod_{j=1}^{m} g_{U_j}(t_j) + g_L(t) \left( \frac{d g_{U_j}(t_j)}{d t_k} \prod_{j=1}^{m} g_{U_j}(t_j) \right) + \lambda_k \left( \frac{d g_{U_j}(t_j)}{d t_k} + \lambda_0 \right) \frac{\partial g_L(t)}{\partial t_k} \right) \right].
\]

Inserting here the formulas (8) leads to

\[
\sum_{x \geq 0} f_s(x) x_k t_k^{x_k - 1} \prod_{i=1, i \neq k}^{m} t_i^{x_i} = T_0 (\lambda_2 T_1 + \lambda_0 T_2 + \lambda (T_3 + T_4)),
\]

where

\[
T_0 = g_s(t) = \sum_{y \geq 0} f_s(y) \prod_{j=1}^{m} t_j^{y_j}, \quad T_1 = g_{U_k}(t_k) = \sum_{z_k=0}^{\infty} f_k(z_k) z_k t_k^{z_k - 1},
\]

\[
T_2 = \frac{\partial g_L(t)}{\partial t_k} = \sum_{z \geq 0} f_0(z) z_k t_k^{z_k - 1} \prod_{j=1, j \neq k}^{m} t_j^{z_j}, \quad T_3 = \frac{\partial g_L(t)}{\partial t_k} \prod_{j=1, j \neq k}^{m} g_{U_j}(t_j) = T_3 \prod_{j=1}^{m} \sum_{u_j=0}^{\infty} f_k(u_j) t_j^{u_j},
\]

\[
T_4 = g_L(t) g_{U_k}(t_k) \prod_{j=1}^{m} g_{U_j}(t_j) = \sum_{z \geq 0} f_0(z) \prod_{j=1}^{m} t_j^{z_j} (\sum_{u_j=0}^{\infty} f_k(u_j) u_j t_k^{u_j - 1}) \prod_{j=1, j \neq k}^{m} \sum_{u_j=0}^{\infty} f_k(u_j) t_j^{u_j}.
\]
We separately evaluate each term of (9). We start with

\[ T_0T_1 = \sum_{y \geq 0} \sum_{z \geq 0} f_k(y) f_k(z) z^t y^t \prod_{l=1}^{m} t_l^{y_l} \]

where we change the variable \( x_k = y_k + z_k \rightarrow y_k = x_k - z_k \), we interchange the sums and obtain

\[ T_0T_1 = \sum_{y \geq 0} \sum_{z \geq 0} f_k(z) f_k(y) f_k(y_1, ..., y_{k-1}, x_k - z_k, y_{k+1}, ..., y_m) z^t y^t \prod_{l=1}^{m} t_l^{y_l} \]

Similarly, changing variables \( x_j = y_j + z_j \), \( \forall j \), and interchanging the sums, yields

\[ T_0T_2 = \sum_{x \geq 0} \sum_{z \geq 0} f_k(x - z) f_k(z) z^t x^t \prod_{l=1}^{m} t_l^{x_l} \]

Also,

\[ T_3 + T_4 = \sum_{x \geq 0} f_0(z) \left( z^t \prod_{l=1}^{m} u_l^{y_l+1} \right) \sum_{u \geq 0} f_k(u) u^t \prod_{l=1}^{m} f_l(u_l) t_l^{u_{l+1}} \]

from where

\[ T_0(T_3 + T_4) = \sum_{y \geq 0} f_k(y) f_0(z) \left( \prod_{l=1}^{m} f_l(u_l) t_l^{y_{l+1}+1} \right) \sum_{u \geq 0} f_k(u) u^t \prod_{l=1}^{m} f_l(u_l) t_l^{u_{l+1}} \]

where we change variables \( v_j = u_j + z_j \) and \( x_j = v_j + y_j = u_j + z_j + y_j \), \( \forall j \), which gives

\[ T_0(T_3 + T_4) = \sum_{u \geq 0} f_k(x - v) f_0(v - u) v^t \prod_{l=1}^{m} f_l(u_l) t_l^{v_{l+1}} \]

Inserting all these formulas into (9) and identifying the coefficients of \( t_k^{v_k+1} \prod_{l=1}^{m} t_l^{y_l} \) yields

\[ x_k f_k(x) = \lambda_k \sum_{z \geq 0} z^k f_k(z) f_k(x_1, ..., x_{k-1}, x_k - z_k, x_{k+1}, ..., x_m) + \lambda_n \sum_{y \geq 0} y^k f_k(y) f_0(v - u) f_k(x - v), \]

from where, if \( x_k \geq 1 \), we easily obtain formula (6).
Particular case. To see how the recursion presented in Proposition 2.2 works, we illustrate it on the particular case \( m = 3 \). The evaluation of \( f_S \) starts with formula (7), which gives \( f_S(0,0,0) \); using then (6), we calculate \( f_S(1,0,0) \) by taking \( k = 1 \), \( f_S(0,1,0) \) by taking \( k = 2 \) and \( f_S(0,0,1) \) by taking \( k = 3 \). For example, we shall use

\[
\nu_f(0,0,0) = \left[ \nu_f(1) + \lambda \nu_f(1,0,0) + \lambda \nu_f(0,1,0) + \lambda \nu_f(0,0,1) \right] \nu_f(0,0,0),
\]

etc. Next step consists of calculating \( f_S(1,1,0) \), \( f_S(1,0,1) \), \( f_S(0,1,1) \) followed by the evaluation of \( f_S(1,1,1) \), and so on.

Remark. If, in particular, we take \( m = 2 \), formula (6) reduces to the particular formulas obtained by Jin and Ren in [5] for their model B with the trivariate Poisson as counting distribution. Moreover, using a similar reasoning as in the above proof, the recursion from Proposition 2.2 can be extended after some tedious calculation to a more general (than the multivariate Poisson) counting distribution, built like the one in the general model B of [5] (which is in fact our model \( 1 \) for \( m = 2 \)); more precisely, the counting distribution results by taking \( N_j = Z_j + Z, 0 \leq j \leq m, \) with independent \( Z_j \)'s, also independent of \( Z \), all of them \( (Z \) included) being distributed in the so-called Panjer class, which consists of the Poisson, binomial and negative binomial distributions (for more details on the Panjer class of distributions see [4]). In conclusion, Theorem 2.2 in [5] can be extended from the bivariate case considered in that work to \( m \geq 3 \).

REFERENCES


Received January 12, 2016