D. Kurzmitteilungen

OLE HESSELAGER, Copenhagen

Recursions for a class of compound Lagrangian distributions

1 Introduction

We consider compound random variables

$$X = \sum_{i=1}^{N} Y_i , \tag{1.1}$$

representing the aggregate claims amount. The severities Y_1, Y_2, ... are assumed to be non-negative integer-valued independent and identically distributed (iid) random variables which are independent of the number N of claims.
For a discrete random variable Z we denote by

$$f_Z(z) = \mathsf{P}\,(Z = z), \qquad z = 0, 1, \dots ,$$

the probability function (pf), and the probability generating function (pgf) is denoted by

$$\varphi_Z(u) = \mathsf{E}\, u^Z = \sum_{z=0}^{\infty} f_Z(z) u^z . \tag{1.2}$$

We also remind that the pgf of a compound random variable (1.1) is given by

$$\varphi_X(u) = \varphi_N(\varphi_Y(u)) . \tag{1.3}$$

2 Lagrangian distributions and a recursive algorithm

A simple probabilistic description of the evolution of an infectious disease (or the spread of a fire) specifies that a diseased person will infect a random number of individuals, and the numbers of infections caused by different individuals are mutually independent and identically distributed. We are interested in the total number N of individuals who are eventually infected, or in fire insurance, the total number of risk units destroyed by the fire. The distribution of N is determined by the distribution of the number of individuals infected by one diseased person. With M being the number of infections caused by the first diseased, and N_i being the total number of individuals who will eventually catch the disease via the ith individual, including the ith individual himself, we may write

$$N = 1 + \sum_{i=1}^{M} N_i . \tag{2.1}$$

It follows from the assumptions above that N_1, N_2, ... are independent of M and iid with the same distribution as N, whence, evaluating the pgf on both sides of (2.1) yields

$$\varphi_N(u) = u\varphi_M(\varphi_N(u)) . \tag{2.2}$$

We observe that any counting distribution with pgf φ_M defines a new counting distribution with pgf φ_N via (2.2), and the class of such counting distributions is known as the *Basic Lagrangian* distributions (BL) (see Johnson et al., 1992, p. 97).
For the class of BL distributions determined by (2.2) it holds that $f_N(0) = 0$, which most easily is seen from (2.1), since $N \geq 1$, and also that $f_N(1) > 0$ when $f_M(0) > 0$. The first three moments $\nu_1 = \mathsf{E}\, N$ and $\nu_i = \mathsf{E}\,(N - \nu_1)^i$, $i = 2, 3$, can be expressed in terms of the corresponding moments $\mu_1 = \mathsf{E}\, M$ and $\mu_i = \mathsf{E}\,(M - \mu_1)^i$, $i = 2, 3$, of M as

$$\begin{aligned} \nu_1 &= 1/(1 - \mu_1) \\ \nu_2 &= \mu_2 \nu_1^3 \\ \nu_3 &= \mu_3 \nu_1^4 + 3\mu_2^2 \nu_1^5 \end{aligned}$$

when $\mu_1 < 1$, and the pf can in general be expressed as

$$f_N(x) = \frac{1}{x!} \left[\frac{\partial^{x-1}}{\partial z^{x-1}} (\varphi_M(z))^x \right]_{z=0} . \tag{2.3}$$

Some simple examples of basic BL distributions are given below.

(a) The *Borel distribution*. For $M \sim$ Poisson(λ) and $\varphi_M(u) = e^{\lambda(u-1)}$ we obtain the Borel distribution with pf

$$f_N(x) = \frac{(x\lambda)^{x-1}}{x!} e^{-x\lambda}, \qquad x \geq 1.$$

(b) The *Consul distribution*. For $M \sim$ Binomial(n, p) and $\varphi_M(u) = (pu + 1 - p)^n$ we obtain the Consul distribution with pf

$$f_N(x) = \frac{1}{x} \binom{nx}{x-1} \left(\frac{p}{1-p} \right)^{x-1} (1-p)^{nx}, \qquad x \geq 1.$$

(c) The geometric distribution on $\{1, 2, \ldots\}$ corresponds to the special case where $M \sim$ Binomial($1, p$).

(d) For $M \sim$ Negative Binomial(α, q) with pgf $\varphi_M(u) = \left(\frac{1-q}{1-q^u} \right)^{\alpha}$ we obtain

$$f_N(x) = \frac{\Gamma(\alpha(x+1)-1)}{\Gamma(\alpha x)\, x!} q^{x-1} (1-q)^{\alpha x}, \qquad x \geq 1.$$

A *general Lagrangian distribution* (GL) is obtained by compounding a BL distribution, and a notable example is the generalized Poisson distribution, which is the distribution of a Poisson sum of Borel distributed variables. This distribution has been considered in the actuarial literature by Goovaerts & Kaas (1991) who derived a recursion for the distribution of the total claims amount (1.1), and by Sharif & Panjer (1995) who improved this recursion. Also Ambagaspitiya & Balakrishnan (1995), who investigated the tail behavior of compound generalized Poisson distributions and gave an alternative algorithm for calculating the compound distribution. If a recursion is available for the compound BL distribution, it is also possible to calculate the distribution of aggregate claims for a great variety of compound GL distributions by a two-step procedure, where the second step involves a standard recursion with the BL distribution as a severity distribution. This was used by Goovaerts & Kaas (1991) and by Sharif & Panjer (1995) for the generalized Poisson distribution, where the second step involves the Panjer recursion for compound Poisson distributions. In relation to recursive calculation of aggregate claims distributions, the passage from

BL distributions to GL distributions does not bring about any essential new structure, and we focus here attention on the BL distributions.
Sharif (1995) has derived recursions for compound BL distributions for the case where the distribution of M belongs to Sundt's (1992) class of counting distributions, which in particular contains the familiar (a, b)-class with pf satisfying the condition

$$f_M(m) = \left(a + \frac{b}{m}\right) f_M(m-1)\,, \qquad m \geq 1\,. \tag{2.4}$$

The condition (2.4) is fulfilled by the Poisson distribution ($a = 0$), the Binomial distribution ($a < 0$) and the Negative Binomial distribution ($0 < a < 1$).
The results of Sharif (1995) were obtained by setting up differential equations for the pgf and identifying coefficients in the corresponding power series. In the following we point out that by considering the shifted version of the BL distributions, obtained by shifting the BL distributions one step to the left such that the support becomes $x = 0, 1, \ldots$, one may obtain jointly a set of simple recursions for the compound distributions corresponding to the shifted as well as the unshifted BL distributions. Thus, let $\widetilde{N} = N - 1$ with pgf

$$\varphi_{\widetilde{N}}(u) = \varphi_N(u)/u\,, \tag{2.5}$$

which by (2.2) is determined by the relation

$$\varphi_{\widetilde{N}}(u) = \varphi_M(u\varphi_{\widetilde{N}}(u))\,. \tag{2.6}$$

We introduce the compound variables

$$X = \sum_{i=1}^{N} Y_i\,, \qquad \widetilde{X} = \sum_{i=1}^{\widetilde{N}} Y_i\,.$$

Because $N = \widetilde{N} + 1$ we have that X is distributed as $\widetilde{X} + Y$, where Y is independent of $\widetilde{X}$ with pf f_Y, such that

$$\varphi_X(u) = \varphi_Y(u)\varphi_{\widetilde{X}}(u)\,. \tag{2.7}$$

Since $\varphi_{\widetilde{X}}(u) = \varphi_{\widetilde{N}}(\varphi_Y(u))$ according to (1.3), we also observe from (2.6) that

$$\varphi_{\widetilde{X}}(u) = \varphi_M(\varphi_Y(u)\varphi_{\widetilde{X}}(u)) \overset{(2.7)}{=} \varphi_M(\varphi_X(u)) \,. \tag{2.8}$$

Equations (2.8), (2.7) form the basis for a class of recursions for calculating jointly the pf's $f_X(x)$ and $f_{\widetilde{X}}(x)$. By comparing (2.8) with (1.3) we observe that the distribution of $\widetilde{X}$ is cast as an aggregate claims distribution with a counting distribution given by the variable M, and a severity distribution which is the (compound) distribution of X. When M has a distribution for which a simple recursive formula for the compound distribution is already available, we may then use (2.8) to write down a recursive formula for the pf $f_{\widetilde{X}}$ involving f_X as a severity distribution, and (2.7) shows that the pf f_X can be calculated by use of the convolution formula when the pf $f_{\widetilde{X}}$ is known. Together, this yields a method for calculating recursively the two pf's $(f_{\widetilde{X}}, f_X)$ jointly. Below we work out the details when the distribution of M is assumed to belong to the (a, b)-class (2.4), but first we need to determine the starting values $(f_{\widetilde{X}}(0), f_X(0))$. By letting $u = 0$ in (2.8) and (2.7) we have that

$$\begin{aligned} f_{\widetilde{X}}(0) &= \varphi_M(f_X(0)) \,, \\ f_X(0) &= f_Y(0) f_{\widetilde{X}}(0) \,. \end{aligned} \tag{2.9}$$

Whence,

$$f_Y(0) = 0 \;\Rightarrow\; f_X(0) = 0 \,, \quad f_{\widetilde{X}}(0) = \varphi_M(0) = f_M(0) \,. \tag{2.10}$$

In the case $f_Y(0) > 0$ it follows from (2.9) that $f_X(0)$ is determined by the equation

$$x = h(x) \,, \quad h(x) = f_Y(0)\varphi_M(x) \,,$$

and it is shown in the appendix that the iterated sequence $x_i = h(x_{i-1})$, for arbitrary starting value $x_0 \in [0, 1]$, converges monotonically towards the unique solution $f_X(0)$ to this equation. Thus, the initial values in the case $f_Y(0) > 0$ are given by

$$\begin{aligned} f_X(0) &= x_\infty = \lim_{i\to\infty} x_i \,, \quad x_i = f_Y(0)\varphi_M(x_{i-1}) \\ f_{\widetilde{X}}(0) &= \varphi_M(x_\infty) \,. \end{aligned} \tag{2.11}$$

The examples (a)–(d) in Section 2 are those where the distribution of M belongs to the (a, b)-class defined by (2.4). In this case, the corresponding compound distribution satisfies Panjer's (1981) recursion, such that (2.8) gives that

$$f_X(x) = \sum_{y=0}^{x} \left(a + \frac{by}{x}\right) f_{\widetilde{X}}(y) f_X(x-y), \tag{2.12}$$

and (2.7) gives that

$$f_{\widetilde{X}}(x) = \sum_{y=0}^{x} f_Y(y) f_X(x-y). \tag{2.13}$$

We note that $f_X(x)$ appears on the right-hand side of (2.12) in the term corresponding to $y = 0$ and via $f_{\widetilde{X}}(x)$ also in the term corresponding to $y = x$. By separating out these terms we obtain the recursive formula

$$f_X(x) = \frac{1}{1 - (2a+b) f_X(0)} \left\{ \sum_{y=1}^{x-1} \left(a + \frac{by}{x}\right) f_{\widetilde{X}}(y) f_X(x-y) \right.$$
$$\left. + (a+b) f_X(0) \sum_{y=1}^{x} f_Y(y) f_X(x-y) \right\}, \tag{2.14}$$

$$f_{\widetilde{X}}(x) = \sum_{y=0}^{x} f_Y(y) f_X(x-y). \tag{2.15}$$

To determine the starting values from (2.11) in the case $f_Y(0) > 0$ we need the pgf $\varphi_M(u)$ which for the (a, b)-class is given by

$$\varphi_M(u) = \begin{cases} e^{b(u-1)}, & a = 0 \ \text{(Poisson)} \\ \left(\frac{1-a}{1-au}\right)^{1+b/a} & a \neq 0 \ \text{(Binomial, Negative Binomial)} \end{cases}. \tag{2.16}$$

In other situations, e.g. when the distribution of M belongs to Sundt's (1992) class of counting distributions, one obtains in a similar manner recursions jointly for the compound shifted and unshifted BL distributions by using the relevant recursive formula in the place of the Panjer recursion (2.12).

In general, we observe from (2.8), (2.7) that the computational effort involved with calculating the compound shifted BL distributions recursively is of the same order as the recursion available for calculating the distribution of a compound sum with counting variable M. In particular, the order of the recursion (2.14), (2.15) is that of the Panjer recursion, which is $O(x^2)$, meaning that the number of computations needed to calculate the values $(f_X(z), f_{\tilde{X}}(z))$ for $z = 0, \ldots, x$ increases as x^2.

References

Ambagaspitiya, R.S. & Balakrishnan, N. (1995). On the Compound Generalized Poisson Distributions. *ASTIN Bulletin*, **24**, p. 255–263.

Goovaerts, M.J. & Kaas, R. (1991). Evaluating Compound Generalized Poisson Distributions Recursively. *ASTIN Bulletin*, **21**, p. 193–198.

Johnson, N.L., Kotz, S. & Kemp, A.W. (1992). *Univariate Discrete Distributions*. 2nd. edition. Wiley, New York.

Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, **12**, p. 22–26.

Sharif, A.H. (1995). Generalized Recursions for Total Claims Distributions. PhD. thesis, University of Waterloo, Canada.

Sharif, A.H. & Panjer, H.H. (1995). An improved Recursion for the Compound Generalized Poisson Distribution. *Mitteilungen der Schweiz. Vereinigung der Versicherungsmathematiker*. Heft **1/1995**, p. 93–98.

Sundt, B. (1992). On Some Extensions of Panjer's Class of Counting Distributions. *ASTIN Bulletin*, **22**, p. 61–80.

Appendix

The function $h(x) = f_Y(0)\varphi_M(x)$ is increasing and convex on $[0,1]$ with $h(0) = f_Y(0)f_M(0) \geq 0$ and $h(1) = f_Y(0) \leq 1$. We therefore conclude that the equation $x = h(x)$ has a unique solution $f_X(0)$ in the interval $[0,1]$, and that $h(x) \geq x$ for $x \leq f_X(0)$ and $h(x) \leq x$ for $x \geq f_X(0)$.
Consider the iterated sequence $x_i = h(x_{i-1})$ with starting value $x_0 < f_X(0)$. When $x_{i-1} \leq f_X(0)$, because h is increasing,

$$x_i = h(x_{i-1}) \leq h(f_X(0)) = f_X(0) ,$$

such that the values x_i are all bounded by $f_X(0)$ when $x_0 \leq f_X(0)$. Furthermore

$$x_i = h(x_{i-1}) \geq x_{i-1} ,$$

because $h(x) \geq x$ for $x \leq f_X(0)$. Whence, the sequence x_i is monotonically increasing and bounded by $f_X(0)$, such that $x_\infty = \lim_{i\to\infty} x_i$ is well-defined, and since $h(x_\infty) = x_\infty$ it follows that $x_\infty = f_X(0)$. Analogously it is seen that the sequence x_i is monotonically decreasing with limiting value $f_X(0)$ for arbitrary starting value $x_0 \geq f_X(0)$.

Ole Hesselager
Laboratory of Actuarial Mathematics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark