RECURSIONS FOR CONVOLUTIONS OF ARITHMETIC DISTRIBUTIONS

BY NELSON DE PRIL*

Instituut voor Actuariële Wetenschappen, K.U. Leuven, Belgium

ABSTRACT

A simple recursion for the n-fold convolution of an arithmetic distribution with itself is developed and its relation to Panjer's algorithm for compound distributions is shown.

KEYWORDS

Recursion, convolution, arithmetic distribution.

1. INTRODUCTION

Let $X_1, X_2, \ldots, X_n$ be n mutually independent random variables with a common arithmetic distribution. By a proper rescaling we may assume these variables to be defined on the integers. We denote their common distribution by $p_X(x)$. First we will focus on the case in which the X_i are defined on the non-negative integers with $p_X(0) > 0$. In chapter 4 the more general case will be discussed. Let

(1) $$G_X(u) = \sum_{x=0}^{\infty} p_X(x) u^x$$

be the probability generating function of X.

The distribution of the sum

(2) $$S = X_1 + X_2 + \cdots + X_n$$

is given by the n-fold convolution

(3) $$p_S(s) = p_X^{*n}(s)$$

and the generating function is

(4) $$G_S(u) = [G_X(u)]^n.$$

The usual method to calculate $p_X^{*n}(s)$ requires $n-1$ successive applications of the formula

(5) $$p_X^{*(k+1)}(y) = \sum_{x=0}^{y} p_X(x) p_X^{*k}(y-x), \qquad k = 1, 2, \ldots, n-1, \quad y = 0, 1, \ldots, s$$

* The author is grateful to the editor and the anonymous referees for their helpful suggestions. He thanks especially one of them for his detailed comments which led to an improved presentation and extended results.

and is thus very time-consuming for large values of n, especially if X takes on more than a few different values.

In this paper a simple recursion is developed which gives the n-fold convolution directly in terms of $p_X(x)$. This permits to reduce the number of required computations in a substantial way.

2. THE RECURSION

THEOREM 1. *Let the X_i be defined on the non-negative integers and $p_X(0) > 0$. Then the following recursion holds*:

$$p_S(0) = [p_X(0)]^n \tag{6a}$$

$$p_S(s) = \frac{1}{p_X(0)} \sum_{x=1}^{s} \left(\frac{n+1}{s}x - 1\right) p_X(x) p_S(s-x), \qquad s = 1, 2, \ldots. \tag{6b}$$

PROOF. Formula (6a) is trivial and can be obtained by putting $u = 0$ in (4). To prove (6b) we can start from

$$G_X(u) G'_S(u) = n G'_X(u) G_S(u). \tag{7}$$

According to Leibnitz's formula taking the derivative of order $s-1$ of both sides of (7) gives:

$$G_X(u) G_S^{(s)}(u) + \sum_{x=1}^{s-1} \binom{s-1}{x} G_X^{(x)}(u) G_S^{(s-x)}(u)$$
$$= n \sum_{x=0}^{s-1} \binom{s-1}{x} G_X^{(x+1)}(u) G_S^{(s-1-x)}(u)$$

and setting $u = 0$ leads to formula (6b).

As suggested by one of the referees a very elegant alternative derivation of (6b) consists in the use of some conditional expectation. This approach was also used by BÜHLMANN and GERBER (1980) in a proof of the algorithm for the compound Poisson distribution.*

Let X_{n+1} be an auxiliary random variable with the same distribution as $X_1, X_2, \ldots, X_n$, but independent of these variables. Then we have

$$E\left[\frac{n+1}{s} X_{n+1} - 1 \,\middle|\, S + X_{n+1} = s\right] = 0 \tag{8}$$

giving

$$\sum_{x=0}^{s} \left(\frac{n+1}{s}x - 1\right) p_X(x) p_S(s-x) = 0$$

from which (6b) follows immediately.

This proof is more direct but the choice of the appropriate conditional expectation requires some *a priori* knowledge of the formula to be obtained.

* In several recent papers the recursion for the compound Poisson distribution is attributed to ADELSON (1966). However an earlier reference is KATTI and GURLAND (1958) as mentioned by SHUMWAY and GURLAND (1960).

We remark that the recursion (6) already exists in the form of an algorithm for computing a power series raised to a power, see e.g. KNUTH (1969, pp. 444–446). Knuth refers to HENRICI (1956), who attributes the result to a communication by J. C. P. Miller.

3. INTERPRETATION AND RELATION TO PANJER'S ALGORITHM

Let K be a random variable defined on the non-negative integers with distribution $p_K(k)$ and generating function $G_K(u)$.

Let $Y_1, Y_2, \ldots$ be a sequence of independent and identically distributed random variables defined on the positive integers with common distribution $p_Y(y)$ and generating function $G_Y(u)$. Assume that the Y_i's are also independent of K and consider

$$R = Y_1 + Y_2 + \cdots + Y_K \tag{9}$$

with the convention that $R = 0$ if $K = 0$.

The distribution of this random sum is

$$p_R(0) = p_K(0) \tag{10a}$$

$$p_R(r) = \sum_{k=1}^{r} p_K(k) p_Y^{*k}(r), \qquad r = 1, 2, \ldots \tag{10b}$$

and the generating function is

$$G_R(u) = G_K[G_Y(u)]. \tag{11}$$

The next theorem shows that for appropriate choices of the distribution of K and Y the sum S can be replaced by the random sum R.

THEOREM 2. *Let*

$$p_K(y) = \binom{n}{k} [1 - p_X(0)]^k [p_X(0)]^{n-k}, \qquad k = 0, 1, \ldots, n \tag{12}$$

and

$$p_Y(y) = \frac{p_X(y)}{1 - p_X(0)}, \qquad y = 1, 2, \ldots \tag{13}$$

then the distributions of S and R are identical.

PROOF. From $G_K(u) = [p_X(0) + (1 - p_X(0))u]^n$ and

$$G_Y(u) = \frac{G_X(u) - p_X(0)}{1 - p_X(0)}$$

it follows immediately that $G_R(u) = G_S(u)$.

The idea behind this theorem is that the n variables in the sum S, of which many may take on the value zero, is replaced by a sum of random number of strictly positive variables. The number of positive terms is binomially distributed with parameters n and $1-p_X(0)$ and the distribution of S can be seen as a compound binomial distribution

(14a)
$$p_S(0)=[p_X(0)]^n$$

(14b)
$$p_S(s)=\sum_{k=1}^{s}\binom{n}{k}[1-p_X(0)]^k[p_X(0)]^{n-k}p_Y^{*k}(s), \qquad s=1,2,\ldots.$$

From this result—which can be verified without calculations—it follows that the recursion (6) can be derived as a special case of the algorithm of PANJER (1981) for the compound binomial distribution. We remark that this fact can also implicitely be found in JEWELL and SUNDT (1981, p. 227) where they state that when the portfolio is homogeneous, the binomial approximation is exact. However, they do not mention that in this case the aggregate claims distribution is just the n-fold convolution of the common claims distribution of the single policies, and that thus the algorithm can be used for the evaluation of the n-fold convolution of these distributions.

4. GENERALIZATION

So far we considered only the case in which the X_i are defined on the non-negative integers and where $p_X(0)>0$. The adaptation of the recursion (6) to the general case is given in the following theorem.

THEOREM 3. *Let the X_i be defined on the integers and let $m=\min\{x: p_X(x)>0\}$, with $m>-\infty$. Then the following recursion holds:*

(15a)
$$p_S(mn)=[p_X(m)]^n$$

(15b)
$$p_S(s)=\frac{1}{p_X(m)}\sum_{x=1}^{s-mn}\left(\frac{n+1}{s-mn}x-1\right)p_X(x+m)p_S(s-x),$$
$$s=mn+1, mn+2,\ldots.$$

PROOF. Let

(16)
$$Z_i=X_i-m$$

(17)
$$T=\sum_{i=1}^{n}Z_i=S-mn$$

then

(18)
$$p_Z(z)=p_X(z+m)$$

(19)
$$p_T(t)=p_S(t+mn).$$

https://doi.org/10.2143/AST.15.2.2015024 Published online by Cambridge University Press

Since $\Pr(Z_i \geqslant 0) = 1$ and $p_Z(0) = p_X(m) > 0$ Theorem 1 can be applied in terms of the variables Z_i and T, so that

$$p_T(0) = [p_Z(0)]^n$$

$$p_T(t) = \frac{1}{p_Z(0)} \sum_{x=1}^{t} \left(\frac{n+1}{t} x - 1\right) p_Z(x) p_T(t-x), \qquad t = 1, 2, \ldots.$$

After insertion of (18) and (19) and introduction of $s = t + mn$ the result (15) is obtained. It is clear that for $m = 0$ we get Theorem 1 as a special case.

REFERENCES

ADELSON, R. M. (1966) Compound Poisson distributions. *Operations Research Quarterly* **17**, 73–75.

BÜHLMANN, H. and GERBER, H. U. (1980) Discussion of "The aggregate claims distribution and stop-loss reinsurance" by H. H. Panjer. *Transactions of the Society of Actuaries* **32**, 537–538.

HENRICI, P. (1956) Automatic computations with power series, *Journal of the Association for Computing Machinery* **3**, 10–15.

JEWELL, W. S. and SUNDT, B. (1981) Improved approximations for the distribution of a heterogeneous risk portfolio. *Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker* 1981, 221–240.

KATTI, S. K. and GURLAND, J. (1958) Some families of contagious distributions. Air Force Technical Note, Air Force Office of Scientific Research Contract AF 49(638)-43, Technical Note No. TN 59-4 AD 208184.

KNUTH, D. E. (1969) *The Art of Computer Programming, Vol. 2, Seminumerical Algorithms.* Addison-Wesley Series in Computer Science and Information Processing.

PANJER, H. H. (1981) Recursive evaluation of a family of compound distributions. *Astin Bulletin* **12**, 22–26.

SHUMWAY, R. and GURLAND, J. (1960) A fitting procedure for some generalized Poisson distributions. *Skandinavisk Aktuarietidskrift* **43**, 87–108.