Probability in the Engineering and Informational Sciences, **27**, 2013, 319–331.
doi:10.1017/S0269964813000053

EXPANSIONS FOR MOMENTS OF COMPOUND POISSON DISTRIBUTIONS

S. Nadarajah
School of Mathematics
University of Manchester
Manchester M13 9PL, UK
E-mail: saraless.nadarajah@manchester.ac.uk

C.S. Withers
Applied Mathematics Group
Industrial Research Limited
Lower Hutt, New Zealand
E-mail: c.withers@irl.cir.nz

S.A.A. Bakar
Institute of Mathematical Sciences
University of Malaya
50603 Kuala Lumpur, Malaysia
E-mail: saab@um.edu.my

Expansions for moments of $\overline{X}$, the mean of a random sample of size n, are given for both the univariate and multivariate cases. The coefficients of these expansions are simply Bell polynomials. An application is given for the compound Poisson variable S_N, where $S_n = n\overline{X}$ and N is a Poisson random variable independent of $X_1, X_2, \ldots$, yielding expansions that are computationally more efficient than the Panjer recursion formula and Grubbström and Tang's formula.

1. INTRODUCTION AND SUMMARY

The compound Poisson model is one of the most popular models in probability, statistics, operations research, and many applied areas (social sciences, economics, management science, electrical and electronic engineering, industrial engineering, biology, etc.). Its classical form is: if $X_1, X_2, \ldots$ is a random sample and $N \sim$ Poisson (λ) independently then

$$S_N = \sum_{i=1}^{N} X_i, \tag{1.1}$$

is the compound Poisson random variable. It is of interest to know the distribution of S_N, in particular its moments.

 0269-9648/13 $25.00

Usually, the distribution of a random variable can be described well by its first four moments. However, there are many situations in insurance and economics that require moments of orders higher than four. We mention:

- Taleb [11] suggests using moments of order higher than four to measure the risk of an option. For example, the fifth moment is suggested as the asymmetry sensitivity of the fourth one. The seventh moment is suggested as the sign of the convexity change as the underlying asset moves up or down.
- Avramidis and Matzinger [2] show that an estimator for pricing American options can be improved using moments of order higher than four.

For some other examples, we refer the readers to Coën, Racicot and Théoret [3].

The aim of this note is to provide general and accessible formulas for moments of (1.1) for both the univariate and multivariate cases. These formulas are given in terms of Bell polynomials. In-built routines for Bell polynomials are available in most computer algebra packages. For example, see BellY in Mathematica and IncompleteBellPoly in Matlab. So, the formulas given will be accessible to most practitioners.

Let $\overline{X} = n^{-1}S_n = n^{-1}\sum_{i=1}^n X_i$ be the mean of a random sample $X_1, \ldots, X_n$ from a distribution with finite moments and cumulants

$$m_r = m_r(X) = \mathbb{E}\left(X^r\right), \quad \mu_r = \mu_r(X) = \mathbb{E}\left(X - m_1\right)^r, \quad \kappa_r = \kappa_r(X),$$

where the cumulants $\{\kappa_r\}$ of X are defined by

$$\sum_{r=1}^{\infty} \kappa_r t^r / r! = \log\left\{\mathbb{E}\left[\exp(tX)\right]\right\}, \tag{1.2}$$

see (Kendall and Stuart [7], Section 3.12). The problem of obtaining expansions in powers of n^{-1} for $m_r(\overline{X})$ and $\mu_r(\overline{X})$ is a very old one. By a laborious method a solution was given by Tchouproff [12]. He showed that these moments are polynomials in n^{-1} of the form

$$m_r\left(\overline{X}\right) = \sum_{i=0}^{r-1} n^{-i} m_{ri}, \tag{1.3}$$

$$\mu_r\left(\overline{X}\right) = \sum_{i=r/2}^{r-1} n^{-i} \mu_{ri}, \tag{1.4}$$

for $r \geq 1$, where $\sum_{i=a}^{b}$ sums over integers i such that $a \leq i \leq b$. He gave the coefficients m_{ri} for $2 \leq r \leq 4$ and μ_{ri} for $2 \leq r \leq 8$ in Eq. (11), page 151 and Eq. (26), page 155.

The results of this note are organized as follows. In Section 2, we prove that the general formulas for the coefficients in (1.3) and (1.4) are given by

$$m_{ri} = B_{r,r-i}(\boldsymbol{\kappa}) = m_{ri}(\boldsymbol{\kappa}), \qquad \mu_{ri} = r! B_{i,r-i}\left(\boldsymbol{\kappa}'\right)/i! = \mu_{ri}(\boldsymbol{\kappa}), \tag{1.5}$$

where

$$\boldsymbol{\kappa} = \left(\kappa_1, \kappa_2, \ldots\right), \quad \boldsymbol{\kappa}' = \left(\kappa_1', \kappa_2', \ldots\right), \quad \kappa_i' = \kappa_{i+1}/(i+1), \tag{1.6}$$

and for $\mathbf{x} = (x_1, x_2, \ldots)$, $B_{rk}(\mathbf{x})$ is defined by

$$\left(\sum_{r=1}^{\infty} x_r t^r / r!\right)^k / k! = \sum_{r=k}^{\infty} B_{rk}(\mathbf{x}) t^r / r!, \tag{1.7}$$

where $B_{rk}(\mathbf{x})$ is called the *partial exponential Bell polynomial* and is tabled on page 307 of Comtet [4] for $r \leq 12$. These results have been known before in a more general setting of infinitely divisible distributions, where expressions for moments in terms of the Bell polynomials in cumulants were given, which were then written as moments of the Lévy measure. However, this is the first time such results have been derived for a compound sum.

In Section 3, we illustrate how the multivariate analogs of (1.3) and (1.4) immediately follow. We also give alternative formulas in terms of $\mathbf{m} = (m_1, m_2, \ldots)$.

In Section 4, we provide applications of these results for the compound Poisson variable in (1.1). We show that formulas based on (1.3) and (1.4) are computationally more efficient than at least two other ways to compute moments of compound sums.

The original contributions of this note are: (1) expansions for moments and central moments of univariate and multivariate compound sums with terms of the expansions expressed in terms of Bell polynomials; (2) application for the compound Poisson distribution; (3) formulas computationally more efficient than Panjer recursion formula and Grubbström and Tang's formula for moments of compound sums. Although the expansions in (1) are special cases of earlier results, it is the first time such expansions have been obtained for compound sums.

2. EXPANSIONS FOR THE MOMENTS OF THE MEAN

Theorem 2.1 proves (1.3) and (1.4) with general formulas for the coefficients m_{ri} and μ_{ri}. Some alternative formulas for the coefficients, m_{ri} and μ_{ri}, are given in Theorem 2.2.

THEOREM 2.1: *Suppose $m_1 = 0$. Then, (1.3) and (1.4) hold with the coefficients, m_{ri} and μ_{ri}, given by (1.5).*

PROOF: From (1.7),

$$\exp\left(\sum_{r=1}^{\infty} x_r t^r / r!\right) = \sum_{r=0}^{\infty} Y_r(\mathbf{x}) t^r / r!,$$

where

$$Y_r(\mathbf{x}) = \sum_{k=0}^{r} B_{rk}(\mathbf{x}),$$

is called the *complete exponential Bell polynomial*. From the definition (1.2),

$$\sum_{r=1}^{\infty} \kappa_r t^r / r! = \log\left\{\mathbb{E}\left[\exp(tX)\right]\right\} = \log\left[\sum_{r=0}^{\infty} m_r t^r / r!\right],$$

so $m_r = Y_r(\boldsymbol{\kappa})$ as noted on page 160 of Comtet [4].

Now

$$\kappa_r\left(S_n\right) = \kappa_r\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \kappa_r\left(X_i\right) = n\kappa_r,$$

and $B_{rk}(n\boldsymbol{\kappa}) = n^k B_{rk}(\boldsymbol{\kappa})$, so

$$m_r\left(S_n\right) = Y_r\left(n\boldsymbol{\kappa}\right) = \sum_{k=0}^{r} B_{rk}\left(n\boldsymbol{\kappa}\right) = \sum_{k=0}^{r} n^k B_{rk}(\boldsymbol{\kappa}).$$

https://doi.org/10.1017/S0269964813000053 Published online by Cambridge University Press

Also

$$B_{r0}(\mathbf{x}) = \delta_{r0} = \begin{cases} 1, & \text{if } r = 0, \\ 0, & \text{if } r \neq 0, \end{cases}$$

so

$$\begin{aligned} m_r\left(\overline{X}\right) &= n^{-r} m_r\left(S_n\right), \\ &= n^{-r} \sum_{k=0}^{r} n^k B_{rk}(\boldsymbol{\kappa}), \\ &= n^{-r} \sum_{i=0}^{r} n^{r-i} B_{r,r-i}(\boldsymbol{\kappa}), \\ &= \sum_{i=0}^{r} n^{-i} B_{r,r-i}(\boldsymbol{\kappa}), \\ &= \sum_{i=0}^{r-1} n^{-i} B_{r,r-i}(\boldsymbol{\kappa}). \end{aligned}$$

This proves (1.3).

Since $m_1 = 0, m_r = \mu_r$. By equation [3l'] of (Comtet [4], p. 136),

$$B_{rk}(\boldsymbol{\kappa})/r! = B_{r-k,k}\left(\boldsymbol{\kappa}'\right)/(r-k)!, \tag{2.1}$$

for $\boldsymbol{\kappa}'$ of (1.6). Since $B_{rk}(\boldsymbol{\kappa}') = 0$ for $k > r$, (1.3) implies

$$\mu_r\left(\overline{X}\right) = \sum_{i=0}^{r-1} n^{-i} B_{r,r-i}\left(\boldsymbol{\kappa}\right) = r! \sum_{i=0}^{r-1} n^{-i} B_{i,r-i}\left(\boldsymbol{\kappa}^{'}\right)/i! = r! \sum_{i=r/2}^{r-1} n^{-i} B_{i,r-i}\left(\boldsymbol{\kappa}^{'}\right)/i!. \tag{2.2}$$

This proves (1.4). ∎

Putting $n = 1$ in (2.2) gives μ_r in terms of the cumulants. This holds regardless of whether in fact $m_1 = 0$, since a change in location does not affect $\{\mu_r\}$.

THEOREM 2.2: *Suppose $m_1 = 0$. Then,*

$$m_{ri} = \sum_{j=r-i}^{r} S_j^{(r-i)} B_{rj}(\mathbf{m}), \tag{2.3}$$

for $r \geq 1$ and

$$\mu_{ri} = \sum_{j=r-i}^{r/2} S_j^{(r-i)} B_{r-j,j}\left(\boldsymbol{\mu}'\right), \tag{2.4}$$

where $\mu'_j = \mu_{j+1}/(j+1)$ and $S_j^{(k)}$ is the Stirling number of the first kind, tabled on page 833 of Abramowitz and Stegun [1].

PROOF: Take the coefficient of $t^r/r!$ in

$$\mathbb{E}\left[\exp\left(tS_n\right)\right] = (1+T)^n,$$

where

$$T = \sum_{i=1}^{\infty} m_i t^i / i!.$$

That is,

$$\mathbb{E}\left(S_n^r\right) = \sum_{j=1}^{r} B_{rj}(\mathbf{m}) n!/(n-j)!,$$

for $r \geq 1$. Since

$$n!/(n-j)! = \sum_{k=0}^{j} n^k S_j^{(k)},$$

we have (2.3). Using (2.1) gives (2.4). ■

However, these alternative formulas, (2.3) and (2.4), are more complex than those of (1.5).

Example 2.1: Note that

$$m_4\left(\overline{X}\right) = \sum_{i=0}^{3} n^{-i} m_{4i},$$

where $m_{40} = B_{44}(\boldsymbol{\kappa}) = \kappa_1^4$, $m_{41} = B_{43}(\boldsymbol{\kappa}) = 6\kappa_1^2\kappa_2$, $m_{42} = B_{42}(\boldsymbol{\kappa}) = 4\kappa_1\kappa_3 + 3\kappa_2^2$, and $m_{43} = B_{41}(\boldsymbol{\kappa}) = \kappa_4$. Also

$$\mu_4\left(\overline{X}\right) = \sum_{i=2}^{3} n^{-i} \mu_{4i},$$

where $\mu_{42} = B_{22}(\boldsymbol{\kappa}')4!/2! = 3\kappa_2^2$ and $\mu_{43} = B_{31}(\boldsymbol{\kappa}')4!/3! = \kappa_4$.

3. MULTIVARIATE EXTENSIONS

Theorems 3.1 and 3.2 are analogs of Theorem 2.1 for the multivariate case.

THEOREM 3.1: *Suppose* $\mathbf{X}$ *and* $\overline{\mathbf{X}}$ *are p-variate, where* $p > 1$. *Set* $\mathbf{X} = (X_1, \ldots, X_p)$, $m^i = \mathbb{E}[X_i]$, *and*

$$\begin{aligned} m^{1\cdots p} &= m^{1\cdots p}(\mathbf{X}) = \mathbb{E}\left(X_1 \cdots X_p\right), \\ \mu^{1\cdots p} &= \mu^{1\cdots p}(\mathbf{X}) = \mathbb{E}\left[\left(X_1 - m^1\right) \cdots \left(X_p - m^p\right)\right], \\ \kappa^{1\cdots p} &= \kappa\left(X_1, \ldots, X_p\right) = \kappa(\mathbf{X}). \end{aligned}$$

Suppose also $m^i = 0$. *Then, for* $p \geq 1$,

$$m^{1\cdots p}\left(\overline{\mathbf{X}}\right) = \sum_{i=0}^{p-1} n^{-i} m_i^{1\cdots p}(\boldsymbol{\kappa}), \tag{3.1}$$

and

$$\mu^{1\cdots p}\left(\overline{\mathbf{X}}\right) = \sum_{i=p/2}^{p-1} n^{-i} \mu_i^{1\cdots p}(\boldsymbol{\kappa}), \tag{3.2}$$

where $m_i^{1\cdots p} = m_i^{1\cdots p}(\boldsymbol{\kappa})$ *and* $\mu_i^{1\cdots p} = \mu_i^{1\cdots p}(\boldsymbol{\kappa})$ *are* $m_{pi}(\boldsymbol{\kappa})$ *and* $\mu_{pi}(\boldsymbol{\kappa})$ *rewritten in the obvious manner in terms of* $\boldsymbol{\kappa}$, *the set of joint cumulants of* $\mathbf{X}$.

PROOF: Follow the proof of Theorem 2.1 (or equivalently (1.3), (1.4)) with $m_r(\overline{X})$ replaced by $m^{1\cdots p}(\overline{\mathbf{X}})$, m_{ri} replaced by $m_i^{1\cdots p}(\boldsymbol{\kappa})$, $\mu_r(\overline{X})$ replaced by $\mu^{1\cdots p}(\overline{\mathbf{X}})$, and μ_{ri} replaced by $\mu_i^{1\cdots p}(\boldsymbol{\kappa})$. ■

THEOREM 3.2: *Under the conditions of Theorem 3.1, for $r \geq 1$ and $\alpha_1, \ldots, \alpha_r$ in $\{1, \ldots, r\}$,*

$$m^{\alpha_1\cdots\alpha_r}\left(\overline{\mathbf{X}}\right) = \sum_{i=0}^{r-1} n^{-i} m_i^{\alpha_1\cdots\alpha_r},$$

and

$$\mu^{\alpha_1\cdots\alpha_r}\left(\overline{\mathbf{X}}\right) = \sum_{i=r/2}^{r-1} n^{-i} \mu_i^{\alpha_1\cdots\alpha_r}.$$

PROOF: Follow the proof of Theorem 2.1 with $m_r(\overline{X})$ replaced by $m^{\alpha_1\cdots\alpha_r}$, m_{ri} replaced by $m_i^{\alpha_1\cdots\alpha_r}$, $\mu_r(\overline{X})$ replaced by $\mu^{\alpha_1\cdots\alpha_r}$, and μ_{ri} replaced by $\mu_i^{\alpha_1\cdots\alpha_r}$. ■

Example 3.1: From Example 2.1,

$$m^{1\cdots 4}\left(\overline{\mathbf{X}}\right) = \sum_{i=0}^{3} n^{-i} m_i^{1\cdots 4},\ \mu^{1\cdots 4}\left(\overline{\mathbf{X}}\right) = \sum_{i=2}^{3} n^{-i} \mu_i^{1\cdots 4},$$

where

$$\begin{aligned}
m_0^{1\cdots 4} &= \kappa^1 \cdots \kappa^4,\ m_1^{1\cdots 4} = \sum^6 \kappa^1\kappa^2\kappa^{34},\\
m_3^{1\cdots 4} &= \kappa^{1\cdots 4},\\
\mu_2^{1\cdots 4} &= \sum^3 \kappa^{12}\kappa^{34},\ \mu_3^{1\cdots 4} = \kappa^{1\cdots 4},
\end{aligned}$$

where $\sum^6$ and $\sum^3$ denote summations over all permutations of $1\ldots 4$ giving distinct terms. For example, $\sum^3 \kappa^{12}\kappa^{34} = \kappa^{12}\kappa^{34} + \kappa^{13}\kappa^{24} + \kappa^{14}\kappa^{23}$.

Example 3.2: Replacing superscript 4 by 1 in Example 3.1 gives

$$m^{1123}\left(\overline{\mathbf{X}}\right) = \mathbb{E}\left[\left(\overline{X}_1\right)^2 \overline{X}_2\overline{X}_3\right] = \sum_{i=0}^{3} n^{-i} m_i^{1123}$$

and

$$\mu^{1123}\left(\overline{\mathbf{X}}\right) = \mathbb{E}\left[\left(\overline{X}_1 - m^1\right)^2 \left(\overline{X}_2 - m^2\right)\left(\overline{X}_3 - m^3\right)\right] = \sum_{i=2}^{3} n^{-i} \mu_i^{1123},$$

where

$$\begin{aligned}
m_0^{1123} &= \left(\kappa^1\right)^2 \kappa^2\kappa^3,\\
m_1^{1123} &= \left(\kappa^1\right)^2 \kappa^{23} + \kappa^2\kappa^3\kappa^{11} + 2\kappa^1\kappa^2\kappa^{13} + 2\kappa^1\kappa^3\kappa^{12},\\
m_2^{1123} &= 2\kappa^1\kappa^{123} + \kappa^2\kappa^{1123} + \kappa^3\kappa^{112} + \mu_2^{1123},\\
\mu_2^{1123} &= \kappa^{11}\kappa^{23} + 2\kappa^{12}\kappa^{13},\\
m_3^{1123} &= \mu_3^{1123} = \kappa^{1123}.
\end{aligned}$$

Replacing superscript 3 by 2 gives

$$m^{1122}\left(\overline{\mathbf{X}}\right) = \mathbb{E}\left(\overline{X}_1^2\overline{X}_2^2\right) = \sum_{i=0}^{3} n^{-i} m_i^{1122},$$

and

$$\mu^{1122}\left(\overline{\mathbf{X}}\right) = \mathbb{E}\left[\left(\overline{X}_1 - m^1\right)^2\left(\overline{X}_2 - m^2\right)^2\right] = \sum_{i=2}^{3} n^{-i}\mu_i^{1122},$$

where

$$\begin{aligned}
m_0^{1122} &= \left(\kappa^1\right)^2\left(\kappa^2\right)^2, \\
m_1^{1122} &= \left(\kappa^1\right)^2\kappa^{22} + \left(\kappa^2\right)^2\kappa^{11} + 4\kappa^1\kappa^2\kappa^{12}, \\
m_2^{1122} &= 2\kappa^1\kappa^{122} + 2\kappa^2\kappa^{112} + \mu_2^{1122}, \\
\mu_2^{1122} &= \kappa^{11}\kappa^{22} + 2\kappa^{12}\kappa^{12}, \\
m_3^{1122} &= \mu_3^{1122} = \kappa^{1122}.
\end{aligned}$$

4. APPLICATION FOR THE COMPOUND POISSON

Applying Theorems 2.1 and 3.1 to the compound Poisson model, we obtain the following.

THEOREM 4.1: *Consider $X_1, X_2, \ldots$ in $\mathbb{R}$ as in Section 2 with $m_1 = 0$ and $N \sim$ Poisson (λ) independently. For $r \geq 1$,*

$$m_r\left(S_N\right) = \sum_{k=1}^{r}\lambda^k B_{rk}(\mathbf{m}) = \sum_{i=0}^{r-1}\lambda^{r-i} m_{ri}(\mathbf{m}), \tag{4.1}$$

where $m_{ri}(\mathbf{m}) = B_{r,r-i}(\mathbf{m})$ and

$$\mu_r\left(S_N\right) = \sum_{k=1}^{r/2}\lambda^k B_{r-k}\left(\mathbf{m}'\right) r!/(r-k)! = \sum_{i=r/2}^{r-1}\lambda^{r-i}\mu_{ri}(\mathbf{m}), \tag{4.2}$$

where $\mu_{ri}(\mathbf{m}) = B_{i,r-i}(\mathbf{m}')r!/i!$. Similarly, if $\mathbf{X}$ and $\{\mathbf{X}_i\}$ are r-variate with $m^i = \mathbb{E}[X_i] = 0$, we have

$$m^{1\cdots r}\left(\mathbf{S}_N\right) = \sum_{i=0}^{r-1}\lambda^{r-i} m_i^{1\cdots r}(\mathbf{m}), \tag{4.3}$$

$$\mu^{1\cdots r}\left(\mathbf{S}_N\right) = \sum_{i=r/2}^{r-i}\lambda^{r-i}\mu_i^{1\cdots r}(\mathbf{m}), \tag{4.4}$$

where $\mathbf{m}$ is the set of joint non-central moments of $\mathbf{X}$.

https://doi.org/10.1017/S0269964813000053 Published online by Cambridge University Press

PROOF: Since

$$\mathbb{E}\left[\exp\left(tS_n\right)\right] = M^n,$$

where

$$M = \mathbb{E}\left[\exp(tX)\right] = \sum_{r=0}^{\infty} m_r t^r / r!,$$

we have

$$\mathbb{E}\left[\exp\left(tS_N\right)\right] = \sum_{n=0}^{\infty} M^n \frac{\lambda^n \exp(-\lambda)}{n!} = \exp\left(\lambda M - \lambda\right).$$

So,

$$\log\left\{\mathbb{E}\left[\exp\left(tS_N\right)\right]\right\} = \lambda M - \lambda = \lambda \sum_{r=0}^{\infty} m_r t^r / r! - \lambda,$$

and

$$\kappa_r\left(S_N\right) = \lambda m_r.$$

This is exactly analogous to $\kappa_r(S_n) = n\kappa_r$, so (1.3), (1.4) hold with n, $n\overline{X} = S_n$ and $\boldsymbol{\kappa}$ replaced by λ, S_N and $\mathbf{m}$, respectively. That is, (4.1) and (4.2) hold. Also (4.3) and (4.4) are analogous to (3.1) and (3.2), respectively. ■

Example 4.1: By Example 2.1, $\mu_4(S_N) = 3m_2^2\lambda^2 + m_4\lambda$.

Example 4.2: By Example 3.1 if $r = 4$

$$\mu^{1234}\left(\mathbf{S}_N\right) = \left(m^{12}m^{34} + m^{13}m^{24} + m^{14}m^{23}\right)\lambda^2 + m^{1234}\lambda.$$

So,

$$\mu^{1123}\left(\mathbf{S}_N\right) = \left(m^{11}m^{23} + 2m^{12}m^{13}\right)\lambda^2 + m^{1123}\lambda,$$

and

$$\mu^{1122}\left(\mathbf{S}_N\right) = \left(m^{11}m^{22} + 2m^{12}m^{12}\right)\lambda^2 + m^{1122}\lambda.$$

The formulas (4.1) and (4.2) provide ways to calculate the moments and central moments of a compound sum. A traditional way to compute moments of a compound sum is to use the well known Panjer recursion formula (Panjer [8]; Panjer and Willmot [9]):

$$\mathbb{E}\left(S_N^r\right) = \sum_{n=0}^{\infty} n^r g_n, \tag{4.5}$$

where g_n satisfies the recurrence relation

$$g_n = \frac{1}{1 - af_0} \sum_{j=1}^{n} \left(a + \frac{bj}{n}\right) f_j g_{n-j}, \tag{4.6}$$

for some $-b \le a < 1$ with the initial condition

$$g_0 = \mathcal{P}_N\left(f_0\right),$$

where $f_k = \Pr(X_i = k)$ and

$$\mathcal{P}_N(t) = \sum_{j=0}^{\infty} q_j t^j,$$

denotes the probability generating function of N. Substituting (4.6) into (4.5), the recurrence formula for the moments of S_N can be made explicit:

$$\begin{aligned}
\mathbb{E}\left(S_N^r\right) &= \frac{1}{1-af_0}\sum_{n=0}^{\infty} n^r \sum_{j=1}^{n}\left(a+\frac{bj}{n}\right) f_j g_{n-j} \\
&= \frac{1}{1-af_0}\left[a\sum_{n=0}^{\infty}\sum_{j=1}^{n} n^r f_j g_{n-j} + b\sum_{n=0}^{\infty}\sum_{j=1}^{n} j n^{r-1} f_j g_{n-j}\right] \\
&= \frac{1}{1-af_0}\left[a\sum_{n=0}^{\infty}\sum_{j=1}^{n}(n-j+j)^r f_j g_{n-j} + b\sum_{n=0}^{\infty}\sum_{j=1}^{n} j(n-j+j)^{r-1} f_j g_{n-j}\right] \\
&= \frac{1}{1-af_0}\left[a\sum_{n=0}^{\infty}\sum_{j=1}^{n}\sum_{k=0}^{r}\binom{r}{k} j^{r-k}(n-j)^k f_j g_{n-j}\right. \\
&\quad \left. + b\sum_{n=0}^{\infty}\sum_{j=1}^{n}\sum_{k=0}^{r-1}\binom{r-1}{k} j^{r-k}(n-j)^k f_j g_{n-j}\right] \\
&= \frac{1}{1-af_0}\left[a\sum_{k=0}^{r}\binom{r}{k}\sum_{j=1}^{\infty}\sum_{n=j}^{\infty} j^{r-k}(n-j)^k f_j g_{n-j}\right. \\
&\quad \left. + b\sum_{k=0}^{r-1}\binom{r-1}{k}\sum_{j=1}^{\infty}\sum_{n=j}^{\infty} j^{r-k}(n-j)^k f_j g_{n-j}\right] \\
&= \frac{1}{1-af_0}\left[a\sum_{k=0}^{r}\binom{r}{k}\sum_{j=1}^{\infty} j^{r-k} f_j \sum_{m=0}^{\infty} m^k g_m + b\sum_{k=0}^{r-1}\binom{r-1}{k}\sum_{j=1}^{\infty} j^{r-k} f_j \sum_{m=0}^{\infty} m^k g_m\right] \\
&= \frac{1}{1-af_0}\left[a\sum_{k=0}^{r}\binom{r}{k} m_{r-k}\mathbb{E}\left(S_N^k\right) + b\sum_{k=0}^{r-1}\binom{r-1}{k} m_{r-k}\mathbb{E}\left(S_N^k\right)\right], \qquad \textbf{(4.7)}
\end{aligned}$$

where $m_k = \mathbb{E}(X^k)$. The formula in (4.7) is heavily used in actuarial science. If N is a Poisson random variable as in Theorem 4.1 then $a = 0$ and $b = \lambda$, so (4.7) reduces to

$$\mathbb{E}\left(S_N^r\right) = \lambda\sum_{k=0}^{r-1}\binom{r-1}{k} m_{r-k}\mathbb{E}\left(S_N^k\right). \qquad \textbf{(4.8)}$$

For other recursive formulas for computing moments of S_N, we refer the readers to Sections 9.1.2, 9.1.3, and 9.2 (univariate case) and Sections 17.1 and 17.2 (multivariate case) of Sundt and Vernic [10]. We have used Panjer recursion formula because of its widespread use. Embrechts and Frei [5] state "Panjer recursion is arguably the most widely used method to "exactly" evaluate compound distributions".

There have also been some non-recursive formulas for moments of S_N developed in the literature. Of these, the one given by Grubbström and Tang [6] is

$$\mathbb{E}\left(S_N^r\right) = r! \sum_{i=1}^{r} \left\{ \sum_{k=1}^{i} \frac{(-1)^{i-k} m_k(N)}{k!} \sum_{j_1 \geq 1, \ldots, j_k \geq 1, j_1 + \cdots + j_k = i} \sum \cdots \sum \frac{1}{j_1 \cdots j_k} \right\} \times \left\{ \sum_{l_1 \geq 1, \ldots, l_i \geq 1, l_1 + \cdots + l_i = r} \sum \cdots \sum \prod_{j=1}^{i} \frac{m_{l_j}(X)}{l_j!} \right\}, \tag{4.9}$$

see Eq. (26) in Grubbström and Tang [6]. This formula is a finite sum of the product of two terms: the first term is a multiple finite sum, and the second term is a multiple finite sum of a finite product. In fact, (4.9) has

$$\frac{1}{2} \sum_{i=1}^{r} i(i+1)(i+3) = \frac{1}{24} r(r+1)(r+2)(3r+13)$$

finite summations, excluding the finite product. The formulas, (4.1) and (4.8), each has only one finite sum. This is not a concrete mathematical evidence of the fact that (4.1) and (4.8) are simpler than (4.9). But we can at least say that (4.9) appears more complicated (i.e., by visual inspection) than (4.1) and (4.8).

We now compare computational efficiencies of (4.1), (4.8), and (4.9). We compute the central processing unit times taken to calculate $\mathbb{E}(S_N^r)$ for $r = 1, 2, \ldots, 10$. We take X_i to have three possible distributions: the Pareto distribution given by the probability density function

$$f(x) = \frac{\alpha}{(1+x)^{\alpha+1}} \tag{4.10}$$

for $x > 0$ and $\alpha > 0$; the log-normal distribution given by the probability density function

$$f(x) = \frac{1}{x\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(\log x - \mu)^2}{2\sigma^2} \right\} \tag{4.11}$$

for $x > 0$, $-\infty < \mu < \infty$, and $\sigma > 0$; and, the gamma distribution given by the probability density function

$$f(x) = \frac{x^{k-1} \exp(-x/\theta)}{\Gamma(k)\theta^k} \tag{4.12}$$

for $x > 0$, $k > 0$ and $\theta > 0$. These are the three most popular models for insurance claim amounts.

Tables 1–3 compare the relative central processing unit times taken to compute (4.8) and (4.9) for $r = 0, 1, \ldots, 10$ and for the three distributions. The relative central processing unit times are: the central processing unit time taken to compute (4.8) divided by the the central processing unit time taken to compute (4.1); the central processing unit time taken to compute (4.9) divided by the the central processing unit time taken to compute (4.1). The Bell polynomials required for (4.1) were computed using the in-built routine IncompleteBellPoly in Matlab. The formulas, (4.1), (4.8), and (4.9), were implemented in Matlab.

Table 1. Central Processing Unit Times Taken to Compute (4.1), (4.8), and (4.9) for $r = 0, 1, \ldots, 10$ and X_i Having the Pareto Distribution, (4.10), with $\alpha = 3$

r	CPU time for (4.8)/ CPU time for (4.1)	CPU time for (4.9)/ CPU time for (4.1)
0	41.44444	62.52699
1	4920.8	7811.065
2	11883.75	18032.06
3	15628.5	24952.51
4	24024.8	37034.21
5	25865.33	40350.99
6	25442	39731.22
7	25325	39866.01
8	27528	42233.42
9	30613.36	45989.19
10	27515.62	41283.33

Table 2. Relative Central Processing Unit Times Taken to Compute (4.1), (4.8), and (4.9) for $r = 0, 1, \ldots, 10$ and X_i Having the Log-Normal Distribution, (4.11), with $\mu = 3$ and $\sigma = 1$

r	CPU time for (4.8)/ CPU time for (4.1)	CPU time for (4.9)/ CPU time for (4.1)
0	0.1304348	0.2013124
1	44.6	69.90376
2	74.33333	118.5143
3	86.2	129.4462
4	154.6	238.9388
5	174.5714	267.693
6	187.875	284.18
7	199.1	313.9088
8	232.6364	371.8147
9	228.6429	344.1241
10	253.0625	400.749

Table 3. Relative Central Processing Unit Times Taken to Compute (4.1), (4.8), and (4.9) for $r = 0, 1, \ldots, 10$ and X_i Having the Gamma Distribution, (4.12), with $k = 3$ and $\theta = 1$

r	CPU time for (4.8)/ CPU time for (4.1)	CPU time for (4.9)/ CPU time for (4.1)
0	37.94118	59.6819
1	4087.167	6434.526
2	11869	17806.46
3	13729.83	21740.86
4	26681.4	40910.31
5	24767.43	38112.55
6	23608.33	37453.63
7	25653.8	40995.02
8	28434.73	43948.24
9	26448.79	39827.06
10	27720.5	43629.23

The central processing unit times taken to compute (4.1), (4.8), and (4.9) were computed by the Matlab code:

```
%{set parameter values%};

t = cputime;
%{Matlab code for computing (4.1)%};
e = cputime-t;

t = cputime;
%{Matlab code for computing (4.8)%};
e = cputime-t;

t = cputime;
%{Matlab code for computing (4.9)%};
e = cputime-t;
```

Here, cputime is a Matlab function. According to a Matlab manual, cputime "returns the total CPU time (in seconds) used by your Matlab application from the time it was started".

The tables show that (4.1) is the most efficient for all $r = 0, 1, \ldots, 10$ and for the three distributions. The second most efficient for all $r = 0, 1, \ldots, 10$ and for the three distributions is (4.8). The least efficient for all $r = 0, 1, \ldots, 10$ and for the three distributions is (4.9).

The central processing unit times for (4.1) are several orders smaller. Although not evident from the tables, the central processing unit times for (4.1) do not appear to change much with r or the three distributions. The central processing unit times for (4.8) and (4.9) appear to increase with increasing r. This gain in computational time by using (4.1) over (4.8) and (4.9) is very significant. This could be crucial for the many problems (4.8) has been applied to in actuarial science and other areas.

The relative performances of (4.1), (4.8), and (4.9) suggested by Tables 1–3 are not surprising. The formula (4.1) is a finite sum of a well known polynomial for which accurate and efficient algorithms are available. The formula (4.8) is a recursion formula. Each step of recursion involves a finite sum, so one can only expect that it is more expensive than (4.1). The formula (4.9) is a sum of product of terms that are multiple finite sums. So, one can only expect that (4.9) is the most expensive formula. However, we do not have any mathematical proof and our observations are purely empirical. A future work is to give a rigorous mathematical proof of the fact that (4.1) is more efficient than (4.8) and (4.9).

Another advantage of (4.1) over (4.8) is that the former can be used for both X_i discrete and continuous. Panjer recursion formula assumes by definition that X_i are discrete although (4.8) applies for both X_i discrete and continuous. Usually, insurance claim amounts are continuous random variables.

Acknowledgments

The authors would like to thank the Editor, the Associate Editor, and the two referees for careful reading and for comments which greatly improved the paper.

References

1. Abramowitz, M. & Stegun, I.A. (1964). *Handbook of mathematical functions*. Washington, DC: US Department of Commerce, National Bureau of Standards, Applied Mathematics Series, vol. 55.

2. Avramidis, A.N. & Matzinger, H. (2002). Convergence of the stochastic mesh estimator for pricing American options. In Proceedings of the 2002 Winter Simulation Conference, E. Yücesan, C.-H. Chen, J.L. Snowdon, & J.M. Charnes (eds.), pp. 1560–1567.
3. Coën, A., Racicot, F.E., & Théoret, R. (2009). Higher moments as risk instruments to discard errors in variables: the case of the Fama and French model. *Journal of Global Business Administration* 1: 2–22.
4. Comtet, L. (1974). *Advanced combinatorics*. Dordrecht: Reidel.
5. Embrechts, P. & Frei, M. (2009). Panjer recursion versus FFT for compound distributions. *Mathematical Methods of Operations Research* 69: 497–508.
6. Grubbström, R.W. & Tang, O. (2006). The moments and central moments of a compound distribution. *European Journal of Operational Research* 170: 106–119.
7. Kendall, M.G. & Stuart, A. (1969). *The advanced theory of statistics*, vol. 1, 3rd ed. London: Griffin.
8. Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin* 12: 22–26.
9. Panjer, H.H. & Willmot, G.E. (1992). *Insurance risk models* Schaumburg: Society of Actuaries.
10. Sundt, B. & Vernic, R. (2009). *Recursions for convolutions and compound distributions with insurance applications*. Heidelberg: Springer Verlag.
11. Taleb, N. (1997). *Dynamic hedging*. New York: John Wiley and Sons.
12. Tchouproff, A.A. (1918). On the mathematical expectation of the moments of frequency distribution. *Biometrika* 12: 140–169.