B.A.J. **1**, I, 107–124 (1995)

A REVIEW OF PANJER'S RECURSION FORMULA AND ITS APPLICATIONS

BY D. C. M. DICKSON, B.SC., PH.D., F.F.A., A.I.A.A.

ABSTRACT

This paper reviews Panjer's recursion formula for evaluation of compound distributions and illustrates how it can be applied to life and general insurance problems.

KEYWORDS

Recursions; Collective Risk Model; Probability of Ruin; Individual Risk Model

1. INTRODUCTION

Since the publication of Panjer's (1981) celebrated paper on recursive calculation of compound distributions, extensive use of his results has been made by practitioners and researchers alike. Kuon, Radtke & Reich (1993) write ". . . the use of Panjer's algorithm has meanwhile become a widespread standard technique for actuaries". It is, therefore, strange that hardly any reference has been made to Panjer's results in the United Kingdom actuarial literature, although some of these results have been included in the U.K. examination syllabus since 1987.

The purpose of this paper is to review Panjer's recursion formula and to show how it provides a solution to two classical actuarial problems: the calculation of the distribution of aggregate claims from a portfolio and the calculation of the probability of ultimate ruin in the classical risk model. We also show how the formula can be applied to the calculation of the aggregate claims distribution for an individual risk model.

This paper contains no new results. Our main aim in writing the paper is to demonstrate that Panjer's recursion formula is a useful and powerful tool which is simple to apply and which has many applications to insurance problems.

We start by stating Panjer's recursion formula.

2. PANJER'S RECURSION FORMULA

In this section we present Panjer's recursion formula. Proofs of the main results are given in the Appendix.

We consider a random variable S defined as:

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

$$S = \sum_{i=1}^{N} X_i$$

with $S = 0$ when $N = 0$. N is a discrete random variable, distributed on the non-negative integers, and $\{X_i\}_{i=1}^{\infty}$ is a sequence of independent and identically distributed (i.i.d.) random variables. We assume that N is independent of this sequence. For example, we might model the total amount of claims in a year under a motor insurance policy as S, with N representing the number of claims under the policy and X_i representing the amount of the ith claim.

In his paper, Panjer considers the cases when the distribution of X_i is continuous, and when it is a discrete distribution on the positive integers. In the following we restrict our attention to the latter case, as it is of more practical interest.

We now introduce the following notation:

$p_n = \Pr[N = n]$ for $n = 0, 1, 2, \ldots$

$f_k = \Pr[X_i = k]$ for $k = 0, 1, 2, \ldots$

$f_k^{n*} = \Pr[X_1 + X_2 + \ldots + X_n = k]$ for $n = 1, 2, 3, \ldots$ and $k = 0, 1, 2, \ldots$

$g_k = \Pr[S = k]$ for $k = 0, 1, 2, \ldots$.

Panjer's recursion formula applies when the probability function of N satisfies the recursion formula:

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1} \quad \text{for } n = 1, 2, 3, \ldots \tag{2.1}$$

where a and b are constants. Sundt & Jewell (1981) show that the only three distributions which satisfy this recursion formula are the Poisson, the binomial and the negative binomial, of which the geometric distribution is a special case.

Result (*Panjer's Recursion Formula*)

When the probability function of N satisfies formula (2.1), and when $f_0 = 0$:

$$g_0 = p_0$$

and for $k = 1, 2, 3, \ldots$:

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

$$g_k = \sum_{j=1}^{k} \left(a + \frac{bj}{k}\right) f_j g_{k-j}.$$

The beauty of the recursion formula is that it is trivial (for a computer!) to calculate successive values of g_k. Given the values of a and b, and the probability function of X_i, we require the value of g_0 to calculate g_1, then the values of g_0 and g_1 to calculate g_2, and so on.

To appreciate fully how useful Panjer's recursion formula is, we must compare it with the alternative method of calculating g_k. By noting that S can take the value k, $k = 1, 2, 3, \ldots$, if $N = n$ $(\leq k)$ and $X_1 + X_2 + \ldots + X_n = k$, we have:

$$\Pr[S = k] = \sum_{n=1}^{k} \Pr[N = n] \Pr[X_1 + X_2 + \ldots + X_n = k]$$

i.e.:

$$g_k = \sum_{n=1}^{k} p_n f_k^{n*} \quad \text{for } k = 1, 2, 3, \ldots . \tag{2.2}$$

It is possible to compute f_k^{n*} recursively for $n = 2, 3, 4, \ldots$ (see De Pril (1985)) and hence to calculate g_k by formula (2.2). However, this approach is computationally more intensive and hence requires substantially more computer time than calculating g_k by the recursion formula. Calculations that take seconds using the recursion formula can take hours using formula (2.2). We give an illustration of this in the next section.

In the remainder of this paper it is convenient to assume that X_i is distributed on the non-negative integers rather than the positive integers. In these circumstances values of g_k are calculated from:

$$g_0 = \sum_{n=0}^{\infty} p_n f_0^n \tag{2.3}$$

and for $k = 1, 2, 3, \ldots$:

$$g_k = \frac{1}{1 - af_0} \sum_{j=1}^{k} \left(a + \frac{bj}{k}\right) f_j g_{k-j}. \tag{2.4}$$

Also, when $f_0 > 0$ the upper limit of summation in formula (2.2) changes to ∞, since it is possible for the sum of any number of X_is to be equal to k. The formula we use in the next section to calculate f_k^{n*} for $n = 2, 3, 4, \ldots$ and $k = 1, 2, 3, \ldots$ when $f_0 > 0$ is:

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

$$f_k^{n*} = f_0^{-1} \sum_{i=1}^{k} \left(\frac{n+1}{k} i - 1 \right) f_i f_{k-i}^{n*}.$$

(See De Pril (1985) for a proof of this result.)

3. Aggregate Claims Distributions

In this section we interpret the variables introduced in the previous section as follows:

— N is the number of claims arising from an insurance portfolio over a fixed period of time, typically one year;
— X_i is the amount of the ith claim; and
— S is the aggregate claim amount.

Our objective is to use the recursion formula to calculate the distribution of S. This model for aggregate claims is referred to as the collective risk model, and in such a model a Poisson or negative binomial distribution is often used as the claim number distribution.

Given data sets consisting of observed numbers of claims and of observed individual claim amounts, it would be standard practice to fit distributions to these data. It is not our purpose to describe how this is done. Techniques for fitting distributions to insurance data are described by Hogg & Klugman (1984) and Panjer & Willmot (1992). Normally we would use a continuous distribution to fit the claim amounts data. However, in order to apply the recursion formula, we require a discrete distribution for X_i, and so we must replace our fitted continuous distribution by a discrete distribution which has similar features. The similarities can range from matching probabilities over given ranges to matching a given number of moments of the distributions over specified ranges. Discussions of different procedures to discretise a continuous distribution are given by Gerber (1982) and Panjer & Lutek (1983). In our numerical examples below, we apply one of the most straightforward discretisation procedures, referred to by Panjer & Lutek (1983) as crude rounding. This method discretises the continuous distribution on $0, h, 2h, \ldots$, where $h > 0$, and matches cumulative probabilities at a given set of points.

Another important point in the discretisation procedure is the choice of monetary unit. Suppose that individual claim amounts are recorded in units of £1,000 and that the fitted distribution has mean 1 (unit). Then, if that distribution is discretised on the integers (units, not £s), we would not expect the resulting distribution to be a particularly good match to the original distribution. If, however, the distribution is discretised on monetary amounts of £10, that is one-hundredths of the mean of the fitted distribution, then

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

intuitively we would expect the discretised distribution to be a much better match to the original distribution.

Example 1

Let us suppose that the data from a portfolio lead us to choose a Poisson distribution with mean 20 as the claim number distribution, so that the aggregate claims distribution is compound Poisson. Suppose, also, that a lognormal distribution with mean 1 and variance 3 (measuring in suitable monetary units) is fitted as the individual claim amount distribution. To apply the recursion formula to calculate the aggregate claims distribution, let us choose to discretise the lognormal distribution on 20ths of its mean.

Since the constants in formula (2.1) are $a = 0$ and $b = 20$, and as our discretisation procedure gives $f_0 > 0$, formula (2.4) simplifies to:

$$g_k = \frac{20}{k} \sum_{j=1}^{k} j f_j g_{k-j} \quad \text{for } k = 1, 2, 3, \ldots$$

with $g_0 = \exp\{-20(1-f_0)\}$.

Since we have discretised the lognormal distribution on 20ths of its mean (of 1) the computed value g_k gives $\Pr[S = 0.05k]$ for $k = 0, 1, 2, \ldots$. The table below shows some percentiles of the aggregate claims distribution.

x	$\Pr[S \leq x]$
18.5	0.5
20.5	0.6
22.9	0.7
26.1	0.8
31.1	0.9
36.1	0.95
48.7	0.99

As a simple application, we can use the facts that $E(S) = 20$ — approximately, since our discretisation procedure produces a distribution whose mean is approximately equal to that of the original lognormal distribution — and $\Pr[S \leq 31.1] = 0.9$, to see that the premium loading required for this portfolio to ensure that there is a 10% probability that aggregate claims exceed the premium income is 55.5%. Figure 1 shows the probability function for S, clearly illustrating the positive skew of the

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

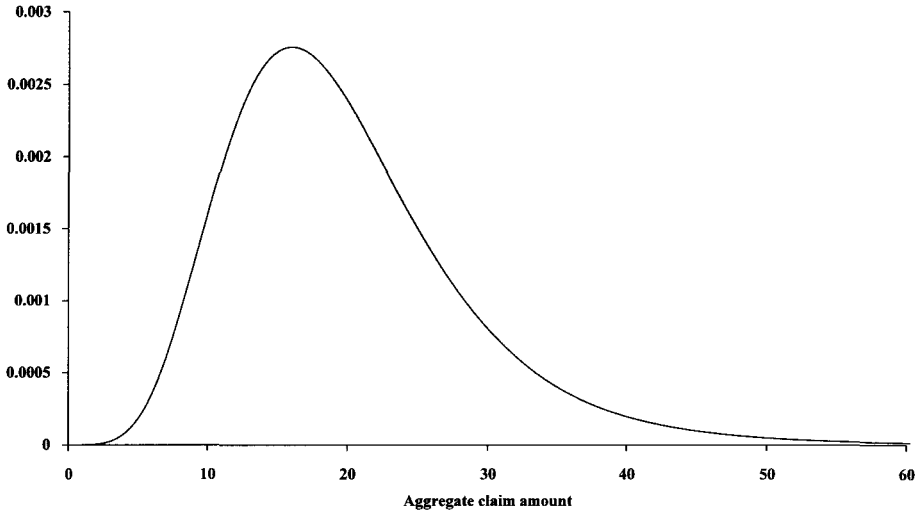

Figure 1. The probability function of aggregate claims from Example 1

distribution. (Note that the large number of points plotted gives the graph the appearance of a density function rather than a probability function.)

To illustrate the computational advantage of the recursion formula over formula (2.2), let us consider some calculation times. Calculation of values of g_k by the recursion formula for values of k from 1 up to 200 took about 1 second. Calculation of the next 200 values took about one more second. To make a comparison of calculation times using formula (2.2) (with $k = \infty$), we have to truncate the summation at some point. We chose to truncate it at ω where ω is the least integer such that:

$$\sum_{n=0}^{\omega} p_n > 0.9999.$$

This ensures that, for any value of k, the difference between the calculated value and the true value of g_k is less than 0.0001. For the calculations described above, the calculation times were around 4 minutes for the first 200 values and 27 minutes for the next 200 values. It is clear that it would require several hours of computer time to produce all the values underlying Figure 1. It took a mere 6 seconds to produce these values using the recursion formula, calculations being done on a 486DX, 33MHz PC.

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

The distribution we have calculated is, of course, only an approximation to the true distribution, as we have replaced a continuous distribution by a discrete one. In general, the approximation should improve as the discretisation procedure becomes more sophisticated. We shall see in the next section that it is possible to assess the accuracy resulting from such an approximation.

One problem that can arise with the recursion formula is that the value of g_0 may be so small that a computer will store it as zero. In such a situation a computer would calculate all values of g_k as zero. The reason for this is that for each value of k, $k = 1, 2, 3, \ldots$, g_k is just some multiple of g_0. This points to one solution to the problem. We simply set g_0 equal to some arbitrary value, for example 1, and calculate values of g_k based on this starting value. Once the required values have been calculated they can be scaled down, for example by multiplying each value n times by $g_0^{1/n}$. It may, in fact, be desirable to scale down calculated values at some intermediate point in the calculation to prevent a numerical overflow in a computer calculation. A full discussion of this problem, including alternative solutions, can be found in Panjer & Willmot (1986, 1992).

4. The Probability of Ultimate Ruin

We now consider an application of the recursion formula when the distribution function of N is geometric with parameter p, where $0 < p < 1$, i.e.:

$$p_n = pq^n \quad \text{for } n = 0, 1, 2, \ldots$$

so that S has a compound geometric distribution.

In this case the constants in formula (2.1) are $a = q$ and $b = 0$, giving:

$$g_0 = p/(1 - qf_0)$$

and for $k = 1, 2, 3, \ldots$:

$$g_k = \frac{q}{1-qf_0} \sum_{j=1}^{k} f_j g_{k-j}. \tag{4.1}$$

If we define:

$$G(k) = \Pr[S \leq k] = \sum_{j=0}^{k} g_j$$

then, by summing over both sides of formula (4.1), we find that, for $k = 1, 2, 3, \ldots$:

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

$$G(k) = g_0 + \frac{q}{1 - qf_0} \sum_{j=1}^{k} f_j G(k - j). \tag{4.2}$$

In other words, when N has a geometric distribution, the recursion formula extends to the distribution function of S. This proves to be extremely useful in the following application.

In this section we use the recursion formula to calculate the probability of ultimate ruin in the classical risk model. Full details of the classical risk model and results in this section are given in the textbook by Bowers *et al.* (1986). It is not our purpose to describe how the formulae below are derived. However, a brief description of the problem is as follows.

In the classical risk model the insurer's surplus at time t is:

$$Z(t) = u + ct - X(t)$$

where u is the insurer's surplus at time 0, c is the insurer's premium income per unit time and $X(t)$ denotes aggregate claims up to time t. The aggregate claims process is a compound Poisson process with Poisson parameter λ, so that λ represents the expected number of claims per unit time. Individual claim amounts have a continuous distribution, with distribution function $P(x)$ and mean μ, and it is assumed that $P(0) = 0$. The insurer's premium income per unit time to cover the risk is $c = (1 + \theta)\lambda\mu$, where θ is referred to as the premium loading factor. The probability of ultimate ruin is denoted by $\psi(u)$ and defined by:

$$\psi(u) = \Pr[Z(t) < 0 \text{ for some } t > 0].$$

The reason why we can apply Panjer's recursion formula to solve for $\psi(u)$ is that, as shown by Bowers *et al.* (1986), we can write:

$$1 - \psi(u) = \Pr[S \leq u]$$

where:

$$S = \sum_{i=1}^{N} X_i \quad (= 0 \text{ if } N = 0).$$

N has a geometric distribution with parameter $1 - \psi(0)$ (which equals $\theta/(1 + \theta)$), and $\{X_i\}_{i=1}^{\infty}$ is a sequence of i.i.d. random variables, independent of N, with distribution function:

$$H(x) = \int_0^x \frac{1 - P(y)}{\mu} dy.$$

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

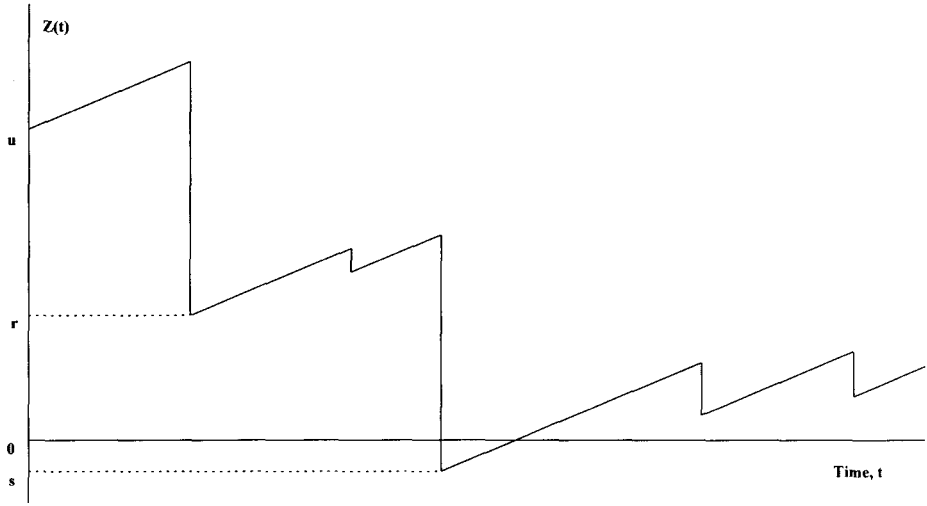

Figure 2. A realisation of the surplus process where ruin occurs at the second increase in the record high of the aggregate loss process

To interpret these variables, we note that the surplus process $\{Z(t)\}_{t\geq 0}$ can be defined in terms of the aggregate loss process $\{L(t)\}_{t\geq 0}$ as:

$$Z(t) = u - L(t)$$

where:

$$L(t) = X(t) - ct$$

so that, for a fixed value of t, $L(t)$ represents the accumulated loss (or gain) up to time t.

Ruin occurs on the first occasion that the aggregate loss exceeds u (if this ever happens). Hence the interpretation of the variables is:

— N denotes the number of occasions on which the aggregate loss process reaches a new record high value;
— X_i represents the amount of the ith *increase* in the record high of the aggregate loss process; and
— S represents the maximum aggregate loss.

Figure 2 shows a realisation of the surplus process where ruin occurs the second time there is a new record high of the aggregate loss process, and Figure 3 shows the corresponding aggregate loss process. In Figure 2, $u - r$ gives the amount of the first increase in the record high of the aggregate loss process, and $r - s$ gives the amount of the second.

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

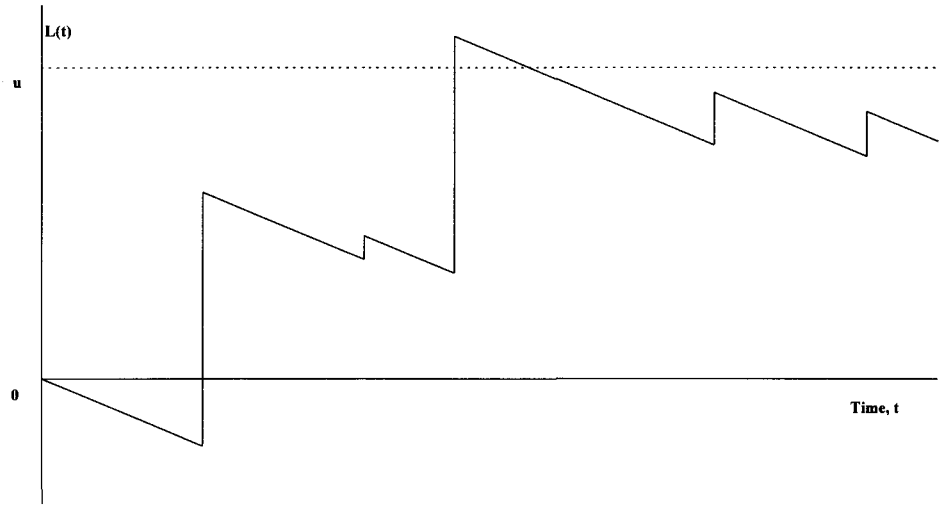

Figure 3. The aggregate loss process corresponding to the realisation of the surplus process in Figure 2

We can apply Panjer's recursion formula, as in the previous section, by replacing $H(x)$ by a discrete distribution. This would give us an approximate value for $1 - \psi(u)$. However, it is actually possible to compute $\psi(u)$ to any number of decimal places by taking the following approach.

Define:

$$H^{n*}(x) = \Pr[X_1 + X_2 + \ldots + X_n \leq x] \quad \text{for } n = 1, 2, 3, \ldots$$

to be the n-fold convolution of the distribution $H(x)$. Then by conditioning on N we can write:

$$1 - \psi(u) = \sum_{n=0}^{\infty} \psi(0)^n (1 - \psi(0)) H^{n*}(u)$$

where $H^{0*}(x)$ is defined to equal 1 for $x \geq 0$, and 0 otherwise.

We can now define two *discrete* distributions with probability functions $l(x)$ and $u(x)$, and distribution functions $L(x)$ and $U(x)$ respectively, as follows:

$$u(x) = H(x + h) - H(x) \qquad \text{for } x = 0, h, 2h, \ldots$$
$$l(x) = H(x) - H(x - h) \qquad \text{for } x = h, 2h, 3h, \ldots.$$

From these definitions we see that:

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

$$L(x) \leq H(x) \leq U(x)$$

and hence:

$$L^{n*}(x) \leq H^{n*}(x) \leq U^{n*}(x)$$

where $L^{n*}(x)$ and $U^{n*}(x)$ are the n-fold convolutions of the distributions $L(x)$ and $U(x)$ respectively. It then follows that:

$$1 - \sum_{n=0}^{\infty} \psi(0)^n(1 - \psi(0))U^{n*}(u) \leq \psi(u) \leq 1 - \sum_{n=0}^{\infty} \psi(0)^n(1 - \psi(0))L^{n*}(u).$$

Thus, it is possible to apply formula (4.2) directly to calculate bounds for $\psi(u)$, since $L(x)$ and $U(x)$ are discrete distributions. As the distributions $L(x)$ and $U(x)$ each have span h, the calculated values will be lower and upper bounds for $\psi(jh)$, $j = 0, 1, 2, \ldots$.

Example 2

Suppose that the individual claim amount distribution is again lognormal with mean 1 and variance 3 (in suitably scaled units) and let the premium loading factor be 20%. The table below shows lower and upper bounds for $\psi(u)$ when $h = 0.05$ and $h = 0.01$, together with the computed value of $\psi(u)$ when $H(x)$ is discretised on intervals of 0.05.

u	Lower bound $h = 0.05$	Lower bound $h = 0.01$	Computed value of $\psi(u)$	Upper bound $h = 0.01$	Upper bound $h = 0.05$
0	0.8261	0.8319	0.8298	0.8333	0.8333
10	0.3218	0.3254	0.3256	0.3269	0.3295
20	0.1553	0.1576	0.1578	0.1586	0.1605
30	0.0795	0.0809	0.0811	0.0815	0.0827
40	0.0424	0.0432	0.0433	0.0436	0.0443
50	0.0234	0.0238	0.0239	0.0241	0.0245
60	0.0133	0.0136	0.0136	0.0137	0.0140
70	0.0078	0.0080	0.0080	0.0080	0.0082
80	0.0047	0.0048	0.0048	0.0049	0.0049

We can see that a reduction in h produces tighter bounds for $\psi(u)$. By choosing smaller values of h we could produce bounds that agree to a given number of decimal places. This has, in fact, happened when $h = 0.01$ and $u = 70$, so we can say that to four decimal places $\psi(70) = 0.0080$. It is clear

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

from the table that the computed values should give a reasonable approximation to the true values, especially for values of $\psi(u)$ less than 0.05, which are the more important ones in practice. A greater degree of accuracy could be obtained by discretising $H(x)$ on a smaller span or by using a more sophisticated discretisation method.

Naturally, this method of lower and upper bounds can be applied in the example of the previous section, although it is not possible to compute the bounds directly as there is no recursion formula for the distribution function of the compound Poisson distribution.

Throughout this section we have assumed that $P(x)$ is a continuous distribution. If $P(x)$ is a discrete distribution, the techniques described above still apply, since $H(x)$ is a continuous distribution even if $P(x)$ is a discrete distribution. The problem of calculating $\psi(u)$ when $P(x)$ is a discrete distribution is discussed by Shiu (1988), who gives an explicit formula for $\psi(u)$.

5. The Individual Risk Model

Consider a fixed number n, of independent policies over a fixed period, say one year. Suppose that the number of claims from each policy will be either zero or one. Then we can model the aggregate claim amount, denoted A, from these policies as:

$$A = Y_1 + Y_2 + \ldots + Y_n$$

where Y_j denotes the amount the insurer pays out under the jth policy. This model is referred to as the individual risk model, as the aggregate claim amount is modelled in terms of payments (zero or otherwise) under individual policies. De Pril (1986) shows how the distribution of A can be calculated recursively. Although this recursion formula is exact, it can be computationally intensive when n is large. Several modifications to this recursion have been suggested; see, for example, De Pril (1988). An alternative to an exact calculation is an approximate calculation using Panjer's recursion formula, as follows.

Let q_j denote the probability that a claim occurs under the jth policy. The amount of that claim is modelled as a random variable with distribution function $P_j(x)$. The distribution of the number of claims from this policy is binomial with parameters 1 and q_j. We can approximate this binomial distribution by a Poisson distribution with parameter q_j, which means that the true compound binomial distribution of Y_j is approximated by a compound Poisson distribution. The reason for making this approximation is that the sum of independent compound Poisson random variables is a compound Poisson random variable (see, for example, Bowers *et al.* (1986), Theorem 11.1), whereas we cannot say what the distribution of the sum of

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

independent compound binomial random variables is. The approximate distribution of A is compound Poisson with Poisson parameter:

$$\lambda = \sum_{j=1}^{n} q_j$$

and individual claim amount distribution:

$$Q(x) = \sum_{j=1}^{n} q_j P_j(x)/\lambda. \tag{5.1}$$

Given this compound Poisson distribution, we can now apply Panjer's recursion formula as in Section 3.

Example 3

A pension scheme provides a death in service benefit to its members. The table below shows the mortality rate and death benefit for members of the scheme at the start of a year.

Mortality rate	Death benefit (£000s)
0.0006	38
0.0006	42
0.0006	45
0.0007	40
0.0007	40
0.0007	42
0.0008	45
0.0018	53
0.0026	48
0.0026	74
0.0030	48
0.0030	55
0.0034	75
0.0038	57
0.0084	78

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

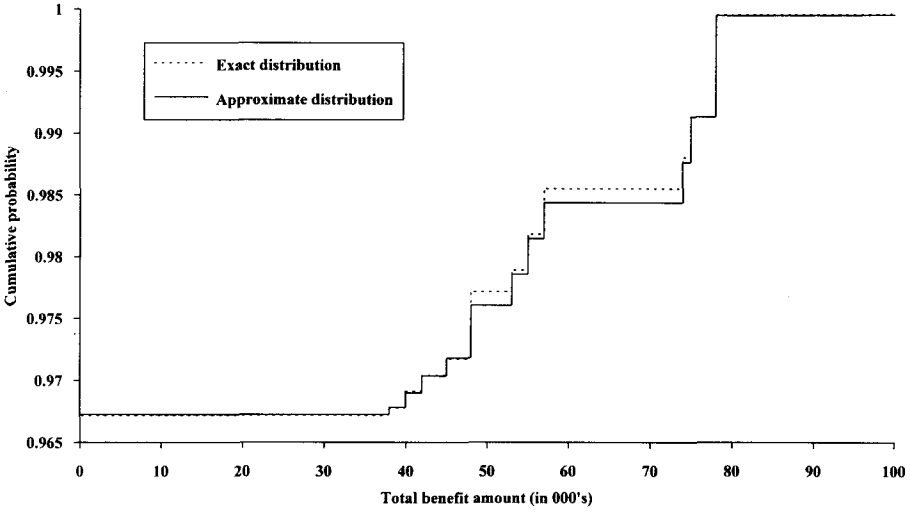

Figure 4. The exact and approximate aggregate claims distributions from Example 3

Using the method described above, we can easily compute the approximate distribution of total death benefits in the year by Panjer's recursion formula. The Poisson parameter is just the sum of the mortality rates, i.e. 0.0333, and the individual claim distribution $Q(x)$ is easily calculated, since the value of each $P_j(x)$ in formula (5.1) is 0 for values of x below the death benefit, and 1 thereafter. Figure 4 shows both exact and approximate values of the distribution function. This graph shows that the compound Poisson distribution provides a good approximation to the true distribution.

Kuon, Reich & Reimers (1987) conclude that this approximation method is sufficiently accurate for practical purposes, and that it has a significant advantage over an exact calculation in terms of computer time required. The accuracy of the approximation is discussed by De Pril & Dhaene (1992), who provide error bounds for the approximation.

6. Concluding Remarks

In the previous three sections we have illustrated applications of Panjer's recursion formula. In each case it is a straightforward exercise to write a computer program to implement the recursion formula. The recursion

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

formula is a simple, but powerful tool for the calculation of compound distributions.

All of the ideas contained in this paper are explored in greater depth in the actuarial literature. Panjer & Willmot's (1992) book deals with the calculation of the aggregate claims distribution for an insurance portfolio, and includes a variety of models which allow recursive calculation of this distribution. This book also contains an extensive bibliography. Panjer & Wang (1993) discuss the stability of the recursion formula. They show that when N has a Poisson or negative binomial distribution the recursion formula is stable. When N has a binomial distribution the recursion formula is unstable, but can still be applied to most practical problems.

Panjer's recursion formula has been generalised for counting distributions of which formula (2.1) is a special case. Sundt & Jewell (1981) consider distributions satisfying:

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1} \quad \text{for } n = 2,3,4,\ldots$$

and Willmot (1988) discusses numerical evaluation of the resulting compound distributions. Schröter (1991) considers distributions satisfying:

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1} + \frac{c}{n} p_{n-2} \quad \text{for } n = 1,2,3,\ldots$$

and Sundt (1992) further generalises to distributions satisfying:

$$p_n = \sum_{i=1}^{k} \left(a_i + \frac{b_i}{n}\right) p_{n-i} \quad \text{for } n = 1,2,3,\ldots$$

with $p_n = 0$ for $n < 0$ in each case. Both authors derive recursion formulae for g_k which generalise Panjer's result.

Other related papers of interest include Bühlmann (1984), who compares the computation of aggregate claims distributions by fast Fourier transforms and by the recursion formula, and Kuon, Reich & Reimers (1987) who compare different methods of computing the aggregate claims distribution for the individual risk model.

References

Bowers, N.L., Gerber, H.U., Hickman, C.J., Jones, D.A. & Nesbitt, C.J. (1986). *Actuarial mathematics*. Society of Actuaries, Itasca, IL.

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

BÜHLMANN, H. (1984). Numerical evaluation of the compound Poisson distribution: recursion or Fast Fourier Transform? *Scandinavian Actuarial Journal*, 116-126.

DE PRIL, N. (1985). Recursions for convolutions of arithmetic distributions. *ASTIN Bulletin*, **15**, 135-139.

DE PRIL, N. (1986). On the exact computation of the aggregate claims distribution in the individual life model. *ASTIN Bulletin*, **16**, 109-112.

DE PRIL, N. (1988). Improved approximations for the aggregate claims distribution of a life insurance portfolio. *Scandinavian Actuarial Journal*, 61-68.

DE PRIL, N. & DHAENE, J. (1992). Error bounds for compound Poisson approximations of the individual risk model. *ASTIN Bulletin*, **22**, 135-148.

GERBER, H.U. (1982). On the numerical evaluation of the distribution of aggregate claims and its stop-loss premiums. *Insurance: Mathematics and Economics*, **1**, 13-18.

HOGG, R.V. & KLUGMAN, S.A. (1984). *Loss distributions*. Wiley, New York.

KUON, S., REICH, A. & REIMERS, L. (1987). Panjer vs Kornya vs De Pril: a comparison from a practical point of view. *ASTIN Bulletin*, **17**, 183-191.

KUON, S., RADTKE, M. & REICH, A. (1993). An appropriate way to switch from the individual risk model to the collective one. *ASTIN Bulletin*, **23**, 23-54.

PANJER, H.H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, **12**, 21-26.

PANJER, H.H. & LUTEK, B.W. (1983). Practical aspects of stop-loss calculations. *Insurance: Mathematics and Economics*, **2**, 159-177.

PANJER, H.H. & WANG, S. (1993). On the stability of recursive formulas. *ASTIN Bulletin*, **23**, 227-258.

PANJER, H.H. & WILLMOT, G.E. (1986). Computational aspects of recursive evaluation of compound distributions. *Insurance: Mathematics and Economics*, **5**, 113-116.

PANJER, H.H. & WILLMOT, G.E. (1992). *Insurance risk models*. Society of Actuaries, Schaumberg, IL.

SCHROTER, K.J. (1991). On a family of counting distributions and recursions for related compound distributions. *Scandinavian Actuarial Journal*, 161-175.

SHIU, E.S.W. (1988). Calculation of the probability of eventual ruin by Beekman's convolution series. *Insurance: Mathematics and Economics*, **7**, 41-47.

SUNDT, B. (1992). On some extensions of Panjer's class of counting distributions. *ASTIN Bulletin*, **22**, 61-80.

SUNDT, B. & JEWELL, W.S. (1981). Further results on recursive evaluation of compound distributions. *ASTIN Bulletin*, **12**, 27-39.

WILLMOT, G.E. (1988). Sundt and Jewell's family of discrete distributions. *ASTIN Bulletin*, **18**, 17-29.

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

APPENDIX

Throughout this appendix we assume that $f_0 > 0$. We prove formulae (2.3) and (2.4) rather than Panjer's recursion formula, which is found by setting $f_0 = 0$ in these formulae. We start by stating and proving two auxiliary results.

Result 1

$$\mathrm{E}\left[X_1 \middle| \sum_{i=1}^{n} X_i = k\right] = \frac{k}{n}.$$

Proof of Result 1

Since $X_1, X_2, \ldots, X_n$ are i.i.d.:

$$\mathrm{E}\left[X_j \middle| \sum_{i=1}^{n} X_i = k\right] = \mathrm{E}\left[X_1 \middle| \sum_{i=1}^{n} X_i = k\right] \quad \text{for } j \neq 1$$

hence:

$$\sum_{j=1}^{n} \mathrm{E}\left[X_j \middle| \sum_{i=1}^{n} X_i = k\right] = \mathrm{E}\left[\sum_{j=1}^{n} X_j \middle| \sum_{i=1}^{n} X_i = k\right] = k$$

and the result follows.

Result 2

For $n = 2, 3, 4, \ldots$:

$$\mathrm{E}\left[X_1 \middle| \sum_{i=1}^{n} X_i = k\right] = \sum_{j=0}^{k} j f_j f_{k-j}^{(n-1)*} / f_k^{n*}.$$

Proof of Result 2

The result comes from noting that for $j = 0, 1, \ldots, k$:

$$\Pr\left[X_1 = j \middle| \sum_{i=1}^{n} X_i = k\right] = \Pr\left[X_1 = j \text{ and } \sum_{i=2}^{n} X_i = k - j\right] \Big/ \Pr\left[\sum_{i=1}^{n} X_i = k\right]$$
$$= f_j f_{k-j}^{(n-1)*} / f_k^{n*}$$

and so:

$$\mathrm{E}\left[\left(X_1 \middle| \sum_{i=1}^{n} X_i = k\right] = \sum_{j=0}^{k} j f_j f_{k-j}^{(n-1)*} / f_k^{n*}.$$

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press

Proof of Formulae (2.3) *and* (2.4)

To prove (2.3), first note that $f_0^{n*} = \Pr[X_1 + X_2 + \ldots + X_n = 0] = f_0^n$ since $\{X_i\}_{i=1}^{\infty}$ is a sequence of i.i.d. random variables. Hence:

$$g_0 = \Pr[S = 0] = p_0 + \sum_{n=1}^{\infty} \Pr[N = n] \Pr[S = 0 \mid N = n]$$

$$= p_0 + \sum_{n=1}^{\infty} p_n f_0^{n*} = \sum_{n=0}^{\infty} p_n f_0^n.$$

For $k = 1, 2, 3, \ldots$:

$$g_k = \sum_{n=1}^{\infty} p_n f_k^{n*}$$

$$= p_1 f_k + \sum_{n=1}^{\infty} p_{n+1} f_k^{(n+1)*}$$

$$= (a + b) p_0 f_k + \sum_{n=1}^{\infty} \left(a + \frac{b}{n+1} \right) p_n f_k^{(n+1)*}.$$

Now $f_k^{(n+1)*} = \sum_{j=0}^{k} f_j f_{k-j}^{n*}$ and by Results 1 and 2:

$$\frac{k}{n+1} f_k^{(n+1)*} = \sum_{j=0}^{k} j f_j f_{k-j}^{n*}.$$

Hence:

$$g_k = (a + b) p_0 f_k + \sum_{n=1}^{\infty} p_n \sum_{j=0}^{k} \left(a + \frac{bj}{k} \right) f_j f_{k-j}^{n*}$$

$$= (a + b) p_0 f_k + \sum_{j=0}^{k} \left(a + \frac{bj}{k} \right) f_j \sum_{n=1}^{\infty} p_n f_{k-j}^{n*}$$

$$= (a + b) p_0 f_k + \sum_{j=0}^{k-1} \left(a + \frac{bj}{k} \right) f_j g_{k-j} + (a + b) f_k (g_0 - p_0)$$

$$\Rightarrow g_k = \frac{1}{1 - a f_0} \sum_{j=1}^{k} \left(a + \frac{bj}{k} \right) f_j g_{k-j}.$$

https://doi.org/10.1017/S1357321700000969 Published online by Cambridge University Press