PANJER vs KORNYA vs DE PRIL: A COMPARISON FROM A PRACTICAL POINT OF VIEW

By S. Kuon, A. Reich and L. Reimers

The Cologne Re, Cologne

ABSTRACT

We compare three modern methods for calculating the aggregate claims distribution with respect to their computation amount and accuracy: Panjer's algorithm, the approximation method of Kornya and the most recent, exact procedure of De Pril. They are compared numerically in the case of actual Life portfolios. The computation amount of De Pril's method is much greater than that of the two others, which do not differ substantially in this respect. The accuracy of Kornya's and Panjer's methods is remarkably high in the examples considered. However, as the accuracy of Kornya's approximation method can be determined easily in advance, this procedure turns out to be the most useful one for the problems arising from practical work.

KEYWORDS

Aggregate claims distribution; computation amount; individual model; collective model.

1. INTRODUCTION

In the years since 1980 a remarkably large number of methods have been made available to insurance mathematicians for calculating the aggregate claims distribution of a portfolio numerically. In the monograph of GERBER (1979) one can find those methods which were used up to 1980 and above all the models to be examined in practice and theory, namely the so-called individual and (due to LUNDBERG (1909)) the collective model of risk theory. Among the methods developed after 1980, of which there is no summarizing description in the international literature, we would like to quote as new ideas or applications the recursive formula of PANJER (1981), the method of Fast Fourier Transform (FFT) of BERTRAM (1981), the approximation method of KORNYA (1983) and the method of DE PRIL (1986). For practitioners and theoreticians, the question arises which of all these methods should be given preference in situations relevant in practice.

With regard to BÜHLMANN (1984), who compared the algorithm of Panjer with the FFT method, it is only necessary to analyse one of these two methods in more detail. We have chosen the Panjer algorithm, because our experience proves that most practitioners choose the Panjer method. Therefore, the aim

of this paper is to compare the methods of Panjer, Kornya and De Pril from a practical point of view.

Such a comparison is feasible only in the case of Life portfolios; i.e. in the language of risk theory in such an individual model, where the claim amount distributions of the individual policies are (individual) two point distributions throughout:

(i) The method of Panjer first of all exactly determines under certain conditions the aggregate claims distribution in a collective model. For a calculation of an individual model, one has to transform this situation into a suitable collective model (due to GERBER (1984) and HIPP (1985) there is an estimate for the error which arises from this transition). The Panjer algorithm, however, then enables the calculation (up to a checkable error term) of the aggregate claims distribution in any individual model.
(ii) The method of Kornya evaluates up to a prescribed accuracy the aggregate claims distribution in an individual model. The original method developed by KORNYA (1983) took only Life portfolios into account, but it can be generalized to an arbitrary individual model, as was shown by HIPP (1986).
(iii) The method of De Pril makes possible the exact calculation of the aggregate claims distribution in the individual model, but only in the case of Life portfolios. At present, a generalization to more general situations is not within sight.

It is therefore precisely in the case of Life portfolios that all three methods can be compared with one another. The fact that in this case the methods of Panjer and Kornya only lead to an approximation is no disadvantage from a practical and theoretical point of view provided that error estimates for the inaccuracies are sufficiently small.

In the following section we present the three methods in a uniform way. In Section 3 we then compare these methods by means of actual portfolios and portfolios derived from these in order to answer the question of which method is the best one for numerical computations. As practitioners, there is no need for us to look at the portfolios given hitherto in actuarial literature, which in general are both theoretical and small.

2. UNIFORM DESCRIPTION OF THE METHODS

For the individual model of a Life portfolio, choose finitely many mortality rates $q_1, \ldots, q_J$ and (up to a fixed monetary unit) finitely many risk sums $i = 1, \ldots, I$, such that each policy of the portfolio has a mortality rate q_{j_0}, $j_0 \leqslant J$, and a risk sum i_0, $i_0 \leqslant I$. Denote by n_{ij} the number of all policies with mortality rate q_j and risk sum i. The distribution function of a policy Y_{ijl} in a cell (i, j), $l = 1, \ldots, n_{ij}$, is

$$F_{ij}(x) = \begin{cases} 1 - q_j, & 0 \leqslant x < i \\ 1, & x \geqslant i. \end{cases}$$

The distribution function

$$F(x) = \text{prob}(S \leqslant x)$$

of the aggregate claim S can be expressed on the (usual) assumption of the independence of all the policies by the convolution formula

$$F(x) = \bigstar_{i,j} F_{ij}^{*n_{ij}}(x)$$

As is well known, for large portfolios it is impossible to carry out these convolutions numerically. Because the support of S is contained in $\mathbb{N}_0$, it is sufficient for a computation of F to calculate

$$f(x) = \text{prob}(S = x), \qquad (x = 0, 1, \ldots).$$

(i) *The method of Panjer.* To apply this method to the given situation one must first of all assign a collective model to the individual model. This can be done in various ways, but we have chosen the most standard procedure. We have chosen a special compound Poisson distribution, which is defined by the equations

$$\lambda = \sum_{i=1}^{I} \sum_{j=1}^{J} q_j n_{ij}$$

$$G(x) = \frac{1}{\lambda} \sum_{i=1}^{I} \sum_{j=1}^{J} q_j n_{ij} F_{ij}^{(0)}(x),$$

where

$$F_{ij}^{(0)}(x) = \begin{cases} 0, & 0 \leqslant x < i \\ 1, & x \geqslant i \end{cases}$$

is independent of j. The density function g of the probability distribution G is given by

$$g(i) = \frac{\lambda_i}{\lambda} \qquad (i = 1, 2, \ldots, I),$$

where

$$\lambda_i = \sum_{j=1}^{J} q_j n_{ij}.$$

The aggregate claims S in the individual model are therefore approximated by an associated collective model

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

where the "claim amounts" X_i are independent and identically distributed with distribution function G. The X_i are also assumed to be independent of the "claim number" N, which follows a Poisson distribution with parameter λ. The density

$$\tilde{f}(x) = \text{prob}(\tilde{S} = x)$$

of the distribution function of $\tilde{S}$ can now be calculated (see PANJER, 1981):

$$\tilde{f}(0) = \exp(-\lambda),$$

and then recursively by

$$\tilde{f}(x) = \frac{\lambda}{x} \sum_{\nu=1}^{\min\{x, I\}} \nu g(\nu) \tilde{f}(x - \nu) \qquad (x = 1, 2, \dots).$$

In case of large λ, $f(0)$ may numerically equal zero. For example, PANJER and WILLMOT (1986) give two different methods (decomposition of the portfolio and exponential scaling) to handle this problem.

For the error

$$\sup_x |F(x) - \tilde{F}(x)|,$$

where F, $\tilde{F}$ denote the corresponding distribution functions, GERBER (1984) and HIPP (1985) have given error estimates.

(ii) *The method of Kornya*. We assume $q_j \leqslant 1/3$, $j = 1, \dots, J$, and define for any $K \in \mathbb{N}$

$$\Delta(K) = \frac{3}{K+1} \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \left(\frac{q_j}{1 - q_j}\right)^{K+1}.$$

In view of

$$\Delta(K) \to 0 \qquad \text{for } K \to \infty$$

one can find to any given accuracy $\varepsilon > 0$ a $K \in \mathbb{N}$ satisfying

$$\exp(\Delta(K)) - 1 \leqslant \varepsilon.$$

For such a fixed K, consider the special polynomial

$$Q^{(K)}(u) = \sum_{k=1}^{K} \frac{(-1)^{k+1}}{k} \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \left[\left(\frac{q_j}{1 - q_j}\right)^k u^{ki} - \left(\frac{q_j}{1 - q_j}\right)^k \right]$$

$$= \sum_{m=0}^{\infty} b_m^{(K)} u^m,$$

where

$$b_0^{(K)} = \sum_{k=1}^{K} \frac{(-1)^k}{k} \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \left(\frac{q_j}{1 - q_j}\right)^k$$

and

$$b_m^{(K)} = \sum_{k=1, k \mid m}^{\min\{K, m\}} \frac{(-1)^{k+1}}{k} \sum_{j=1}^{J} n_{m/k, j} \left(\frac{q_j}{1 - q_j}\right)^k, \qquad m \leqslant IK,$$

denote the non-vanishing Taylor coefficients of $Q^{(K)}$.

The simple recursion formula

$$a_0^{(K)} = \exp b_0^{(K)}$$

$$a_n^{(K)} = \frac{1}{n} \sum_{m=1}^{n} m a_{n-m}^{(K)} b_m^{(K)}, \qquad n \geqslant 1,$$

then leads to the coefficients of the power series of $\exp Q^{(K)}(u)$, i.e.

$$\exp Q^{(K)}(u) = \sum_{n=0}^{\infty} a_n^{(K)} u^n.$$

Finally, if one defines

$$F^{(K)}(x) = \sum_{n=0}^{x} |a_n^{(K)}|,$$

then the result of KORNYA (1983) states that

$$| F(x) - F^{(K)}(x) | \leqslant \exp(\Delta(K)) - 1 \leqslant \varepsilon$$

holds for all $x \leqslant \sum i n_{ij}$.

(iii) *The method of De Pril.* Putting

$$A_{ik} = (-1)^{k+1} i \sum_{j=1}^{J} n_{ij} \left(\frac{q_j}{1 - q_j}\right)^k \qquad (i = 1, \ldots, I, k \geqslant 1),$$

the result of DE PRIL (1986) states that

$$f(0) = \prod_{i=1}^{I} \prod_{j=1}^{J} (1 - q_j)^{n_{ij}}$$

$$f(x) = \frac{1}{x} \sum_{i=1}^{\min(I,x)} \sum_{k=1}^{[x/i]} A_{ik} f(x - ki)$$

holds for all $x \leqslant \sum i n_{ij}$.

3. NUMERICAL COMPARISON

As all of the methods will turn out to be sufficiently precise, we only need to look at the computation amount as an obvious measure of their usefulness. The amount of computation can be quantified as the number of floating point operations, counted separately as bar operations (additions and substractions) and dot operations (multiplications and divisions), or as the computation time (CPU time). We must emphasize that CPU time depends heavily on computer type, programming language and style, and that the turnaround time may be many times longer. The following measurements have been carried out on programs written in VS APL on an IBM 4381 under VM/CMS.

Crucial for the computation amount is not only the size of the portfolio but also the number of values of the aggregate claims distribution to be computed. Given a certain step width Δ (e.g. a fixed number of monetary units) the values

of $F(0)$, $F(\Delta)$, ..., $F(L \cdot \Delta)$ are to be computed for a certain natural number L. In the following examples we have chosen L so that $L \cdot \Delta$ is about the size of $\mu + 3 \cdot \sigma$, where μ and σ denote the mean and standard deviation of the aggregate claims distribution, which can easily be computed beforehand.

In the algorithms of Panjer and Kornya, the computation amount depends on I, J, K (Kornya only) and L, but not on the portfolio size. The number of bar and dot operations (BO resp. DO) can be expressed explicitly. For large L we have, in the case of Panjer's algorithm, approximately $BO \approx DO \approx I \cdot L$; in the case of Kornya's algorithm, approximately $BO \approx DO \approx K \cdot I \cdot L$. De Pril's algorithm needs about $BO \approx L \cdot H_I \cdot (0.5 \cdot L + J - 1) + J$ and $DO \approx L \cdot H_I \cdot (0.5 \cdot L + J + 1.5) + L \cdot (J + 1) + M - 1$, where

$$M = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij}$$

is the number of policies and

$$H_I = \sum_{i=1}^{I} \frac{1}{i} \leqq 1 + \ln I$$

is the ith harmonic number.

The first example derives from a real Life insurance portfolio. It consists of 104,652 risks aged from 15 to 74 with risk sums from DM 2,000 to DM 200,000 in steps of DM 2,000 (that is $I = 100$, $J = 60$). The expected aggregate claim amounts to $\mu = \text{DM}$ 4.537 million, the standard deviation is $\sigma = \text{DM}$ 0.511 million. The aggregate claims distribution is to be computed for a value of $L = 3,000$ which corresponds to an amount of DM 6 million or approximately $\mu + 3 \cdot \sigma$ and comprises 99.6% of its mass.

Algorithm	BO	DO	CPU seconds	Error estimate
Panjer	298,049	307,150	4.200	$2 \cdot 10^{-4}$
Kornya ($K = 5$)	1,530,560	1,378,250	9.748	$3.3 \cdot 10^{-8}$
De Pril	24,121,798	24,443,487	≈ 500	0

Owing to lack of computer capacity, we had to omit the actual computation of $\tilde{F}$ with De Pril's algorithm and could only count the number of operations.

To estimate the accuracy of Kornya's algorithm, we used the formula from Section 2; to estimate the error of the distribution computed according to Panjer, we added the distance between the Kornya and Panjer distribution to the Kornya bound, applying the triangle inequality. According to Section 2 it is possible to estimate the error of the collective model following HIPP (1985), but for this another application of Panjer's algorithm is needed and the computation amount increases considerably. (What is more, Hipp's estimate is rather pessimistic, especially in the case of large portfolios.)

This example demonstrates the usefulness of the above given formulae for the number of FLOPs of Kornya's, Panjer's and De Pril's algorithms. Owing to different program structures, the number of FLOPs per second differs somewhat for the algorithms.

A consideration concerning the choice of step width Δ: halving Δ doubles both L and I. From the formulae given above, it follows as a rule of thumb that all the algorithms need four times the previous computation amount.

We will now examine the performance of DePril's algorithm in comparison with Panjer's and Kornya's if applied to smaller portfolios. To achieve this, we take subportfolios from the example above, consisting of the I youngest age classes and the J smallest sum classes. For Kornya's algorithm we choose $K = 5$; L is again determined from μ and σ.

CASE I. $I = J = 10$; 10,953 risks; $\mu =$ DM 137,346; $\sigma =$ DM 45,861; $L = 250$, 99.99% of mass.

Algorithm	BO	DO	CPU seconds	Error estimate
Panjer	2,304	3,065	0.231	$1.05 \cdot 10^{-4}$
Kornya ($K = 5$)	13,060	11,525	0.327	$5.6 \cdot 10^{-15}$
De Pril	97,334	112,744	4.250	0

CASE II. $I = J = 15$; 27,687 risks; $\mu =$ DM 467,757; $\sigma =$ DM 99,046; $L = 400$, 99.89% of mass.

Algorithm	BO	DO	CPU seconds	Error estimate
Panjer	5,719	6,935	0.386	$1.10 \cdot 10^{-4}$
Kornya ($K = 5$)	31,215	27,625	0.549	$1.3 \cdot 10^{-14}$
De Pril	281,928	319,060	9.779	0

CASE III. $I = J = 20$; 46,698 risks; $\mu =$ DM 945,215; $\sigma =$ DM 156,920; $L = 700$, 99.68% of mass.

Algorithm	BO	DO	CPU seconds	Error estimate
Panjer	13,509	15,630	0.750	$1.14 \cdot 10^{-4}$
Kornya ($K = 5$)	72,120	65,750	1.030	$2.5 \cdot 10^{-14}$
De Pril	923,999	991,118	22.878	0

Our last example is a Life portfolio generated from scratch. It consists of 1,019 risks with sums ranging from DM 2,000 to DM 50,000 (step width is 2,000, i.e. $I = 25$) and ages ranging from 15 to 64 years ($J = 50$). The mortality is 50% of the German mortality table ADSt 60/62 mod. To avoid a voluminous table, the class sizes n_{ij} are defined by the somewhat artificial formula

$$n_{ij} = [0.5 + \max\{0, \ 1.7 \cdot \exp(-0.0005 \cdot (4i^2 + j^2)) + 0.5 \cdot \cos(i + j)\}]$$
$$1 \leqslant i \leqslant 25, \ 1 \leqslant j \leqslant 50.$$

($[x]$ denotes the largest integer less than or equal to x). The aggregate claims distribution was computed for $L = 100$ (approximately $\mu + 3 \cdot \sigma$, 99.5% of mass).

Algorithm	BO	DO	CPU seconds	Error estimate
Panjer	3,349	3,675	0.104	$6.7 \cdot 10^{-4}$
Kornya $(K = 3)$	11,375	4,825	0.158	$2.3 \cdot 10^{-6}$
Kornya $(K = 5)$	16,525	5,150	0.174	$3.4 \cdot 10^{-10}$
De Pril	36,504	43,470	3.180	0

4. CONCLUSIONS

De Pril's algorithm, which gives exact results, is a remarkable progress in theory, but involves much greater computation amount than the two other methods. Panjer's algorithm gives an approximation to the aggregate claims distribution in minimal time. Its accuracy is sufficient for most practical purposes, but it can only be estimated at some additional expense.

Kornya's algorithm needs more computations than Panjer's, but it allows us, given K, to estimate its accuracy, or, if we ask for a specific accuracy, to compute K. In each example the actual CPU time was only slightly higher than that for Panjer's algorithm. The accuracy of Kornya's algorithm, if applied to small portfolios, is so high that it computes practically the exact distribution.

Kornya's algorithm turns out to be a method for computing aggregate claims distributions which is well suited for both small and large Life portfolios.

AUTHORS' NOTE

M. Vandebroek and N. De Pril informed us at proof-stage that they have drawn a similar conclusion, where the exact method of De Pril is taken as a basis for a new approximation procedure. This paper will be published in a special issue of the *Bulletin of the Royal Association of Belgian Actuaries*, dedicated to the 80th birthday of Professor E. Franckx.

REFERENCES

BERTRAM, J. (1981) Numerische Berechnung von Gesamtschadenverteilungen. *Blätter der DGVM* **15**, 175–194.

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

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

GERBER, H. U. (1979) *An Introduction to Mathematical Risk Theory*. Huebner Foundation Monograph, Philadelphia.

GERBER, H. U. (1984) Error bounds for the compound Poisson approximation. *Insurance: Mathematics and Economics* **3**, 191–194.

HIPP, C. (1985) Approximation of aggregate claims distribution by compound Poisson distributions. *Insurance: Mathematics and Economics* **4**, 227–232.

HIPP, C. (1986) Improved approximations for the aggregate claims distribution in the individual model. *ASTIN Bulletin* **16**, 89–100.

KORNYA, P. S. (1983) Distribution of aggregate claims in the individual risk theory model. *Society of Actuaries: Transactions* **35**, 823–858.

LUNDBERG, F. (1909) Zur Theorie der Rückversicherung. *Transactions of the International Congress of Actuaries.*

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

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

S. KUON, A. REICH, L. REIMERS
The Cologne Re, Department for Research and Development, Theodor-Heuss-Ring 11, D-5000 Köln 1, Federal Republic of Germany.