Zeitschrift: Mitteilungen / Schweizerische Aktuarvereinigung = Bulletin / Association Suisse des Actuaires = Bulletin / Swiss Association of Actuaries

Band: - (1996)

Heft: 2

Artikel: An extension of Kornya's method with application to pension funds

Autor: Dufresne, François

DOI: https://doi.org/10.5169/seals-551208

FRANÇOIS DUFRESNE, Lausanne

An extension of Kornya's method with application to pension funds

1 Introduction

Kornya (1983) suggested a method to compute the aggregate claims distribution of a life insurance portfolio. His algorithm gives an approximation as close as desired to the exact distribution under the so called individual model. It can be applied when there is only a single death benefit (no double indemnity or the like). Later on, De Pril (1986) derived an algorithm for the exact calculation of the same distribution. De Pril (1989) modified his algorithm in order to achieve a greater efficiency. Recently, Waldmann (1994) found a way to considerably reduce the number of arithmetic operations of De Pril's exact algorithm.
The aim of this paper is to extend Kornya's method to the case where there are two amounts at risk. Such a situation occurs in the risk analysis of a pension fund when (only) the death and disability benefits of the active members are taken into account. The net amounts at risk are assumed nonnegative. This case can be handled by the algorithms of De Pril (1988) for arbitrary positive claims but our derivation of the algorithm will be along the lines of Kornya and will focus on pension funds applications. An alternative approximation has been proposed by Hipp (1986).

2 The Individual Model

In the individual model of Risk Theory (see, for example, Chapter 2 of Bowers *et al.* (1986)), the aggregate claims random variable S is defined by

$$S = X_1 + X_2 + \cdots + X_m \tag{1}$$

where $X_1, X_2, \ldots, X_m$ are mutually independent random variables. The random variable X_k gives the total claim amount of the insured number k of the portfolio for a given period of time.
In theory, the distribution of S can be obtained recursively by convolution of the distribution of the partial sum $\sum_{j=1}^{k-1} X_j$ with the distribution of X_k,

$k = 2, \ldots, m$. This can be extremely time consuming, even if the X_k's are distributed on the integers.

From now, we consider the risk induced by the death and disability benefits of active members of a pension fund. In this context, it is natural to use the individual model.

Accordingly, the number of active members of the pension fund is denoted by m. For the k-th such member, let d_k and e_k denote the net amounts at risk under the death benefit and disability benefit, respectively, for the considered period of time. (The net amount at risk is the value of the benefit less the corresponding reserve.) The probability that this individual dies in the period of time is q_k and the probability that he or she becomes disabled is i_k. Then, X_k is defined by

$$X_k = \begin{cases} 0 & \text{with probability } p_k \\ d_k & \text{with probability } q_k \\ e_k & \text{with probability } i_k \end{cases} \tag{2}$$

where $p_k = 1 - q_k - i_k$, for $k = 1, 2, \ldots, m$. The distribution of X_k is triatomic if d_k and e_k are not equal and different from zero.

We further assume that d_k and e_k are positive integers[1]. In practice, this would result from an appropriate rounding and change of monetary unit (scale). It follows from this last assumption that S is distributed on $\{0, 1, 2, \ldots\}$. The probability function of S will be denoted by

$$f_x = P[S = x], \qquad x = 0, 1, 2, \ldots \tag{3}$$

Occasionally we will also write $f_S(0)$ instead of f_0. We also assume that $q_k + i_k < 1/2$ for all k. This last assumption will be motivated later.

3 Probability Generating Functions

The probability generating function (p.g.f.) $\varphi_Y(t)$ of a (discrete) random variable Y is defined by

$$\varphi_Y(t) = E[t^Y]. \tag{4}$$

[1] In section 6, we explain how to deal with net amounts at risk which are equal to zero.

The p.g.f. of S will be denoted $\varphi_S(t)$ and for simplicity we shall write $\varphi_k(t)$ for the p.g.f. of X_k. It is well known that the p.g.f. of a sum of independent random variables is the product of the individual probability generating functions.

For the triatomic distribution of X_k, we find that

$$\begin{aligned}\varphi_k(t) &= p_k + q_k t^{d_k} + i_k t^{e_k} \\ &= p_k\left(1 + \frac{q_k}{p_k}t^{d_k} + \frac{i_k}{p_k}t^{e_k}\right) \\ &= p_k(1 + \tilde{q}_k t^{d_k} + \tilde{i}_k t^{e_k}) \qquad (5)\end{aligned}$$

where $\tilde{q}_k = q_k/p_k$ and $\tilde{i}_k = i_k/p_k$.

From (1) and (5), it follows that the p.g.f. of S is

$$\begin{aligned}\varphi_s(t) &= \prod_{k=1}^{m} \varphi_k(t) \\ &= \prod_{k=1}^{m} p_k(1 + \tilde{q}_k t^{d_k} + \tilde{i}_k t^{e_k}) \\ &= f_0 + f_1 t + f_2 t^2 + \cdots + f_x t^x + \cdots + f_L t^L \qquad (6)\end{aligned}$$

where L (a constant) is the maximal aggregate claim and the last equality comes from the definition of the probability generating function. According to (6), if one wants to determine the probability that S equals x, the problem reduces to finding the coefficient of t^x, $x = 0, 1, \ldots, L$.

4 Extension of Kornya's Method

Inspired by Kornya (1983), we rewrite the p.g.f. of S in the following way:

$$\begin{aligned}\varphi_S(t) &= \prod_{k=1}^{m} p_k(1 + \tilde{q}_k t^{d_k} + \tilde{i}_k t^{e_k}) \\ &= \left(\prod_{k=1}^{m} p_k\right) \exp\left(\ln \prod_{k=1}^{m} (1 + \tilde{q}_k t^{d_k} + \tilde{i}_k t^{e_k})\right) \\ &= f_S(0) \cdot \exp\left\{\sum_{k=1}^{m} \ln(1 + \tilde{q}_k t^{d_k} + \tilde{i}_k t^{e_k})\right\}. \qquad (7)\end{aligned}$$

(It is interesting to note that Kornya did not extract $f_S(0)$ in his original paper. This introduces a very small additional error[2] in his method and is the reason for the difference between the approximation of De Pril (1988) and Kornya's approximation.)
Following again Kornya, we use the MacLaurin series expansion of $\ln(1+z)$ and notice that, this time, z is a binomial that we expand according to the binomial formula; it follows that

$$\begin{aligned}\varphi_S(t) &= f_S(0)\cdot\exp\left\{\sum_{k=1}^{m}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}(\tilde{q}_k t^{d_k}+\tilde{i}_k t^{e_k})^n\right\}\\ &= f_S(0)\cdot\exp\left\{\sum_{k=1}^{m}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sum_{l=0}^{n}\binom{n}{l}(\tilde{q}_k t^{d_k})^l(\tilde{i}_k t^{e_k})^{n-l}\right\}.\end{aligned} \tag{8}$$

The expansion of $\ln(1+z)$ is allowed since we assumed that $q_k+i_k<1/2$. We note that the argument of the exponential function in (8) is simply a power series in t. Let $B(t)=\ln\varphi_S(t)$. From (8) it follows that

$$B(t)=\ln f_S(0)+\sum_{k=1}^{m}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sum_{l=0}^{n}\binom{n}{l}(\tilde{q}_k t^{d_k})^l(\tilde{i}_k t^{e_k})^{n-l}. \tag{9}$$

Then we can write

$$\varphi_S(t)=e^{B(t)}. \tag{10}$$

The desired algorithm is a consequence of the following theorem which is essentially due to Euler:

Theorem: *If $A(s)$ and $B(s)$ are power series given by*

$$A(s)=\sum_{i=0}^{\infty}a_i s^i \quad \text{and} \quad B(s)=\sum_{j=0}^{\infty}b_j s^j \tag{11}$$

and if, furthermore, the following relation exists between $A(s)$ and $B(s)$,

$$A(s)=e^{B(s)} \tag{12}$$

[2] In a later stage of the development.

then

$$A(0) = a_0 = \exp(b_0) \tag{13}$$

and

$$a_n = \frac{1}{n} \sum_{k=1}^{n} k b_k a_{n-k}\,, \qquad n = 1, 2, 3, \ldots \tag{14}$$

Proof: For the proof, one considers the identity $A'(s) = B'(s)\,A(s)$ and compares the coefficients of s^{n-1} on both sides.

The algorithm results from the truncation of the infinite power series $B(t)$. Let $B^{(r)}(t)$ be the polynomial obtained by truncating $B(t)$ after the r-th term in the infinite sum:

$$B^{(r)}(t) = \ln f_S(0) + \sum_{k=1}^{m} \sum_{n=1}^{r} \frac{(-1)^{n+1}}{n} \sum_{l=0}^{n} \binom{n}{l} (\tilde{q}_k t^{d_k})^l (\tilde{i}_k t^{e_k})^{n-l}. \tag{15}$$

We say that $B^{(r)}(t)$ is the approximation of order r of $B(t)$.

In summary, to find the distribution of S, one has to ...

1. Choose r, the order of the approximation.
2. Compute $f_S(0) = \prod_{k=1}^{m} p_k$, the initial value.
3. Determine the coefficients of t^k in $B^{(r)}(t)$, say $b_k^{(r)}$.
4. Compute recursively

$$f_x^{(r)} = \frac{1}{x} \sum_{k=1}^{x} k b_k^{(r)} f_{x-k}^{(r)}\,, \qquad \text{for} \quad x = 1, 2, 3, \ldots \tag{16}$$

5 Computational Remarks

For most pension funds, it is useless to classify the risks according to their individual claim distributions: they would be (almost) all different. Of course, this depends on the monetary unit used in the calculations, but in the current practice that would be the case. As a consequence, the preferred way to construct the polynomial $B^{(r)}(t)$ would be given by the following pseudo-code:

For $k = 1$ to m
    For $n = 1$ to r
        For $l = 0$ to n
            Accumulate in position $d_k \cdot l + e_k(n - l)$
            of a one-dimensional array, say B, the quantity
            $\frac{(-1)^{n+1}}{n}\binom{n}{l}\tilde{q}_k^l \cdot \tilde{i}_k^{n-l}$

The outer loop in the preceding pseudo-code means that the polynomial should be constructed while reading the data from the pension fund. The calculation of $f_S(0)$ would be performed at the same time.
It can be shown (see, for example, De Pril (1988)) that, if $q_k + i_k < 1/2$ for all k, then

$$\sum_{x=0}^{\infty} |f_x - f_x^{(r)}| \leq e^{\varepsilon(r)} - 1 \tag{17}$$

where

$$\varepsilon(r) = \frac{1}{r+1} \sum_{k=1}^{m} \frac{p_k}{p_k - q_k - i_k} \left(\frac{q_k + i_k}{p_k} \right)^{r+1}. \tag{18}$$

This last result means that it is possible to determine the order r of the approximation that will provide the desired degree of accuracy.
If the initial value f_0 is very small, the algorithm can be unstable. To avoid such problem, one can use a rescaling strategy as suggested, for example, in Waldmann (1994).

6 Numerical illustration

We will illustrate the method with the data of Held (1982) as they appeared in print[3]. The underlying portfolio consists of 230 active members of a pension fund; like Held, we will call it PK-230.

[3]Dr. Olivier Deprez told the present author that two probabilities in the data of Held were misprinted. In order to let the reader reproduce our results, we use the printed data. Of course, our results are slightly different from those of Held.

To conform to our convention that d_k and e_k are positive integers, one has to set q_k or i_k equal to zero if the corresponding net amounts at risk are zero. Alternatively, one could use the relation $f_S(0) = \exp(b_0)$ where b_0 is the constant term in the right-hand side of (15).

We suppose that a precision of 10^{-6} on the calculated distribution function is desired. The first step consists in choosing the order r of approximation. Table 1 shows the values of $\varepsilon(r)$ computed according to (18). It appears that an approximation of order $r = 5$ is sufficient to satisfy our requirement about the precision.

Table 1 Values of $\varepsilon(r)$ for PK-230

r	$\varepsilon(r)$
1	0.025570
2	9.18495×10^{-4}
3	4.26274×10^{-5}
4	2.22873×10^{-6}
5	1.24694×10^{-7}
6	7.29702×10^{-9}
7	4.41431×10^{-10}
8	2.74164×10^{-11}
9	1.74033×10^{-12}
10	1.12547×10^{-13}
11	7.39707×10^{-15}
12	4.93116×10^{-16}
13	3.32887×10^{-17}
14	2.27249×10^{-18}
15	1.56691×10^{-19}
16	1.09010×10^{-20}
17	7.64491×10^{-22}
18	5.40011×10^{-23}
19	3.83922×10^{-24}
20	2.74544×10^{-25}

The second step indicated at the end of section 4 is the calculation of $f_S(0)$. In fact, we can perform steps 2 and 3 at the same time while reading the data from a computer file (as mentioned in section 5). Table 2 shows the first 31 coefficients of the polynomials $B^{(1)}(t), \ldots, B^{(5)}(t)$. For the illustration, the only polynomial which is needed is $B^{(5)}(t)$; the other polynomials are given to show the speed of convergence of the coefficients.

Table 2 The first 31 coefficients of the polynomials $B^{(r)}(t)$

x	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 5$	x
0	−1.24741404	−1.24741404	−1.24741404	−1.24741404	−1.24741404	0
1	0.17594565	0.17535644	0.17535965	0.17535963	0.17535963	1
2	0.00637411	0.00637411	0.00637411	0.00637411	0.00637411	2
3	0.00587984	0.00585953	0.00585953	0.00585953	0.00585953	3
4	0.05810848	0.05752379	0.05752976	0.05752970	0.05752970	4
5	0.00000000	−0.00001729	−0.00001729	−0.00001729	−0.00001729	5
6	0.05034425	0.05021945	0.05021984	0.05021984	0.05021984	6
7	0.00000000	−0.00168830	−0.00165425	−0.00165477	−0.00165476	7
8	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	8
9	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	9
10	0.05588196	0.05541370	0.05548349	0.05548147	0.05548151	10
11	0.00000000	−0.00041944	−0.00041697	−0.00041698	−0.00041698	11
12	0.05232505	0.05185679	0.05186117	0.05186113	0.05186113	12
13	0.00541659	0.00541659	0.00541659	0.00541374	0.00541385	13
14	0.02555245	0.02551878	0.02551887	0.02551887	0.02551887	14
15	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	15
16	0.01541074	0.01538750	0.01539279	0.01539275	0.01539288	16
17	0.02515861	0.02509621	0.02509641	0.02509641	0.02509641	17
18	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	18
19	0.02058408	0.01927428	0.01929779	0.01929746	0.01929746	19
20	0.09229438	0.09143390	0.09144201	0.09144193	0.09144194	20
21	0.01322054	0.01321352	0.01321353	0.01321345	0.01321345	21
22	0.00888956	0.00888956	0.00888956	0.00888956	0.00888956	22
23	0.01271624	0.01145776	0.01148118	0.01148085	0.01148086	23
24	0.00637896	0.00637896	0.00637896	0.00637896	0.00637896	24
25	0.05172377	0.05124084	0.05124523	0.05124519	0.05124519	25
26	0.01111503	0.01109956	0.01109959	0.01109959	0.01109960	26
27	0.00078069	0.00066100	0.00066144	0.00066144	0.00066144	27
28	0.00000000	−0.00003630	0.00000544	0.00000427	0.00000429	28
29	0.02600034	0.02597711	0.02597716	0.02597716	0.02597716	29
30	0.02471478	0.02457585	0.02457680	0.02457680	0.02457680	30
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 3 Exact and approximation of order $r = 5$ of the distribution function (monetary unit: 1000 CHF)

x	$F(x)$	$F^{(5)}(x)$	$F(x) - F^{(5)}(x)$
0	0.287246646	0.287246646	-3.30×10^{-13}
20	0.417618694	0.417618699	-5.04×10^{-9}
40	0.530181888	0.530181896	-7.78×10^{-9}
60	0.619778881	0.619778895	-1.35×10^{-8}
80	0.681769152	0.681769171	-1.82×10^{-8}
100	0.738048814	0.738048838	-2.39×10^{-8}
120	0.787930329	0.787930359	-3.03×10^{-8}
140	0.820359133	0.820359169	-3.54×10^{-8}
160	0.877666394	0.877666437	-4.22×10^{-8}
180	0.904231654	0.904231701	-4.67×10^{-8}
200	0.925605502	0.925605554	-5.24×10^{-8}
220	0.940434174	0.940434231	-5.73×10^{-8}
240	0.952306117	0.952306178	-6.07×10^{-8}
260	0.962297660	0.962297724	-6.42×10^{-8}
280	0.971451715	0.971451782	-6.67×10^{-8}
300	0.978431097	0.978431166	-6.87×10^{-8}
320	0.983247305	0.983247376	-7.04×10^{-8}
340	0.986882189	0.986882260	-7.17×10^{-8}
360	0.989650826	0.989650899	-7.28×10^{-8}
380	0.991812331	0.991812405	-7.36×10^{-8}
400	0.993392527	0.993392601	-7.43×10^{-8}
420	0.994711650	0.994711725	-7.48×10^{-8}
440	0.996243118	0.996243193	-7.52×10^{-8}
460	0.997122071	0.997122147	-7.56×10^{-8}
480	0.997817798	0.997817874	-7.58×10^{-8}
500	0.998344606	0.998344682	-7.60×10^{-8}

The last step is the calculation of the probability function according to (16). The distribution function can also be computed during this recursion. The "exact" distribution function obtained by direct convolution is presented in

Table 3 with its approximation of order $r = 5$. The errors $F(x) - F^{(5)}(x)$ are also given in this table.[4] It is easily seen that the error is always smaller than the required precision of 10^{-6} on the approximating distribution function.

7 Conclusion

For pension funds, this extension of Kornya's method is much more efficient than brute force convolution. It is very similar to the alternative approach suggested by De Pril on page 23 of his 1989 paper but requires less computer resources. It is also easy to implement.

References

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

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. (1989) The Aggregate Claims Distribution in the Individual Model with Arbitrary Positive Claims. *ASTIN Bulletin* **19**, 9–24

Gerber, H.U. (1979) *An Introduction to Mathematical Risk Theory.* Huebner Foundation for Insurance Education, Philadelphia.

Held, R. (1982) Zur rekursiven Berechnung von Stop Loss-Prämien für Pensionkassen. *MVSVM, Heft 1,* 67–89.

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 the Aggregate Claims in the Individual Risk Theory Model. *Transactions of the Society of Actuaries* **35**, 823–836. Discussion 837–858.

Panjer, H.H. (1981) Recursive Evaluation of a Family of Compound Distributions. *ASTIN Bulletin* **12**, 22–26.

Waldmann, K.-H. (1994) On the Exact Calculation of the Aggregate Claims Distributions in the Individual Life Model. *ASTIN Bulletin* **24**, 89–96.

François Dufresne
Ecole des HEC
University of Lausanne
CH-1015 Lausanne, Switzerland

[4] Note that round-off error is always present in the calculations.

Summary

A simple extension of the method of Kornya is derived. The extended method applies to the convolution of triatomic distributions with nonnegative support while the original method is restricted to diatomic distributions. This way, the algorithm can be applied to the calculation of the distribution of the total claims of a pension fund where only death and disability of active members are considered.

Résumé

On développe une extension de la méthode de Kornya. Cette extension s'applique à la convolution de distributions triatomiques de support non négatif alors que l'application de la méthode originale se restreignait aux distributions diatomiques. Ainsi, l'algorithme peut être appliqué au calcul de la distribution du montant total des sinistres d'une caisse de pension où l'on ne tient compte que des risques de décès et d'invalidité des membres actifs.

Zusammenfassung

Die Methode von Kornya wird verallgemeinert. Die verallgemeinerte Methode kann zur Berechnung der Faltungen von Dreipunktverteilungen mit nichtnegativem Träger angewendet werden, während Kornyas Methode ursprünglich auf Zweipunktverteilungen limitiert ist. Dadurch eignet sich der Algorithmus für die Berechnung des Gesamtschadens einer Pensionskasse, bei der die Risiken Tod und Invalidität der Aktiven betrachtet werden.