Recursions for the individual risk model*

Jan Dhaene† Carmen Ribas‡ Raluca Vernic§

April 14, 2006

Abstract

In the actuarial literature, several exact and approximative recursive methods have been proposed for calculating the distribution of a sum of mutually independent compound Bernoulli distributed random variables. In this paper, we give an overview of these methods. We compare their performance with the straightforward convolution technique by counting the number of dot operations involved in each method. It turns out that in many practicle situations, the recursive methods outperform the convolution method.

1 Introduction

In the individual risk theory model, the aggregate claims of an insurance portfolio over a certain reference period, e.g. 1 year, is modelled as the sum of the claims produced by the individual policies, see e.g. Kaas et al. (2001). Suppose that the claim amounts of the individual policies involved are random variables with a support that is a subset of the non-negative integers and with a strictly positive probability of no claims. Further, assume that the claim amounts of the policies involved are mutually independent. The distribution of the aggregate claims can then be determined by convolution. As this technique is often time-consuming, given the large number of terms in the sum, approximate techniques to calculate quantities such as probabilities and stop-loss premiums have been proposed in the literature, see e.g. Bowers et al. (1997) or Kaas et al. (2001). These techniques are mainly based on fitting moments of the distribution.

Alternatively, several exact and approximate recursive procedures for computing the aggregate claims distribution in the individual risk model have been proposed in the

*Raluca Vernic acknowledges the support from the Dutch Organization for Scientific Research (no. NWO 048.031.2003.001). Jan Dhaene acknowledges the financial support of the Onderzoeksfonds K.U.Leuven (GOA/02: Actuariële, financiële en statistische aspecten van afhankelijkheden in verzekerings- en financiële portefeuilles).

†Katholieke Universiteit Leuven, Belgium and Universiteit van Amsterdam, The Netherlands.

‡University of Barcelona, Spain.

§Ovidius University of Constanta, Romania.

actuarial literature, see De Pril (1988), (1989), Dhaene & De Pril (1994), Dhaene & Sundt (1997, 1998), Dhaene & Vandebroek (1995), Hipp (1985), (1986) and Kornya (1983). Recursions for stop-loss premiums in the individual risk model have been derived by Dhaene, Willmot & Sundt (1999).

In the framework of the individual risk model, the case where the individual claim amounts have only two possible outcomes, 0 and some strictly positive value, has been called the individual life model because of its straightforward application in life insurance. Recursions for this particular case have been derived in De Pril (1986) and Waldmann (1994). A recursion for the number of policies in the portfolio that lead to a strictly positive claim amount has been derived by White & Greville (1959). Most of the recursions that have been derived for the general individual risk model can be considered as generalizations of this early result.

In this paper, we give an overview of the actuarial literature on recursive methods for the computation of aggregate claims distributions within the framework of the individual risk model. Furthermore some new results are obtained in this setting by comparing the different exact and approximative recursions with respect to the straightforward convolution method.

The recursions in this paper can be considered as the equivalent in the individual risk theory model of the many recursions that have been developed for the collective risk model, starting with the seminal papers of Panjer (1981) and Sundt & Jewell (1981). A general framework for recursions of both the individual and collective risk models is introduced and explored in Sundt (1995), (1998), (2000a) and Sundt, Dhaene & De Pril (1998). An extended survey of the literature on recursive evaluation of aggregate claims distributions is given in Sundt (2002).

2 Exact recursions for the individual risk model

2.1 Notations

Consider a portfolio of n policies whose claim amounts are mutually independent. This portfolio is divided into a number of classes by clustering all policies with the same probability of a strictly positive claim amount and with the same severity distribution, as is displayed in the next table. The severity distribution of a policy corresponds to the conditional distribution of the claim amount, given that a claim has occurred.

		claim probability					
		q_1	q_2	$\cdots$	q_j	$\cdots$	q_b
	$g_1(x)$						
	$g_2(x)$				$\cdots$		
claim	$\cdots$						
amount	$g_i(x)$			$\cdots$	n_{ij}		
distribution	$\cdots$						
	$g_a(x)$						

The class (i,j), $i=1,\ldots,a$, $j=1,\ldots,b$, contains all policies with severity distribution g_i and claim probability $q_j = 1-p_j$. The number of policies in class (i,j) is denoted by n_{ij}. It is assumed that for each j, $0<q_j<1$ and that the claim amounts of the individual policies are positive integer representing multiples of some convenient monetary unit, so that for each i, $g_i(x)$ is defined for $x=1,2,\ldots$.

The probability that the aggregate claims S equal s, i.e. $\Pr[S=s]$, is denoted by $p(s)$. For the computation of the probabilities $p(s)$ several computation techniques can be used. We start by presenting the traditional convolution method for the individual risk model.

Let X_{ijk} denote the random claim amount of the k-th policy belonging to the class (i,j), $k=1,...,n_{ij}$. Then the distribution of X_{ijk} can be represented by the matrix

$$\begin{pmatrix} 0 & 1 & 2 & ... & x & ... \\ p_j & q_j\ g_i(1) & q_j\ g_i(2) & ... & q_j\ g_i(x) & ... \end{pmatrix},\ i=1,\ldots,a,\ j=1,\ldots,b, \tag{1}$$

where the first row describes the possible outcomes of the random variable and the second row gives the corresponding probabilities.

The aggregate claims S can be written as

$$S=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} X_{ijk}.$$

Denoting $h_{ij}(x)=\Pr[X_{ijk}=x]$, the probabilities of S follow from

$$p(x)=\underset{i=1}{\overset{a}{*}}\ \underset{j=1}{\overset{b}{*}}\ h_{ij}^{*n_{ij}}(x),$$

where $*$ denotes the convolution operation:

$$(f*g)(x)=\sum_{y=0}^{x} f(x-y)\,g(y), \tag{2}$$

and where $h_{ij}^{*n_{ij}}$ stands for the n_{ij}-fold convolution of h_{ij}.

In the remainder of Section 2, we will consider exact recursive procedures for determining the aggregate claims distribution. In Section 3, we will investigate several approximate recursive methods.

2.2 De Pril's recursion.

One possible way to derive an exact recursion is to introduce coefficients $t(x)$ which are defined by

$$\ln P(u)=\sum_{x=0}^{\infty} t(x)\ u^x, \tag{3}$$

with $P(u)$ being the probability generating function of the aggregate claims S,

$$P(u) = \sum_{s=0}^{\infty} p(s)\ u^s = \prod_{i=1}^{a} \prod_{j=1}^{b} \left[p_j + q_j G_i(u)\right]^{n_{ij}}, \tag{4}$$

and where $G_i(u)$ is the probability generating function of the severity of a particular risk of class (i, j):

$$G_i(u) = \sum_{x=1}^{\infty} g_i(x)\ u^x. \tag{5}$$

Equation (3) gives rise to the following relation:

$$u\ \frac{d}{du} \ln \left(\sum_{x=0}^{\infty} p(x)\ u^x \right) = \sum_{x=0}^{\infty} x\ t(x)\ u^x. \tag{6}$$

Equating the coefficients of the same power of u in (6), we find the following recursion for the probabilities $p(s)$ in terms of the coefficients $t(x)$:

$$\begin{aligned} p(0) &= e^{t(0)}, \\ s\ p(s) &= \sum_{x=1}^{s} x\ t(x)\ p(s-x), \qquad s = 1,\ 2, \ldots. \end{aligned} \tag{7}$$

In this way, the problem of computing the probabilities $p(s)$ can be solved by deriving explicit expressions for the coefficients $t(x)$. In order to derive these expressions, observe that $\ln P(u)$ can be expanded as

$$\begin{aligned} \ln P(u) &= \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \left[\ln p_j + \ln \left(1 + \frac{q_j}{p_j} G_i(u) \right) \right] \\ &= \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \left[\ln p_j + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \left(\frac{q_j}{p_j} \right)^k G_i^k(u) \right], \end{aligned}$$

so that it can be written in the form (3), with coefficients $t(x)$ given by

$$\begin{aligned} t(0) &= \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \ln p_j = \ln p(0), \\ t(x) &= \sum_{i=1}^{a} \sum_{k=1}^{x} A(i,k)\ g_i^{*k}(x), \qquad x = 1,\ 2, \ldots, \end{aligned} \tag{8}$$

where, as before, g_i^{*k} denotes the k-fold convolution of g_i, and the $A(i,k)$ are defined by

$$A(i,k) = \frac{(-1)^{k+1}}{k} \sum_{j=1}^{b} n_{ij}\ \left(\frac{q_j}{p_j} \right)^k. \tag{9}$$

Inserting the $t(x)$-values in (7) leads to an exact recursive formula for the evaluation of the aggregate claims distribution in the individual model:

$$p(s) = \frac{1}{s} \sum_{i=1}^{a} \sum_{k=1}^{s} A(i,k) \sum_{x=k}^{s} x \; g_i^{*k}(x) \; p(s-x), \qquad s = 1, \; 2, \ldots, \tag{10}$$

with initial value $p(0)$ given by

$$p(0) = \prod_{i=1}^{a} \prod_{j=1}^{b} (p_j)^{n_{ij}} \, . \tag{11}$$

This recursion was derived by De Pril (1989) and is now commonly known in actuarial circles as 'De Pril's recursion'. In some sense, it can be considered as the twin brother of Panjer's (1981) recursion in the collective model.

2.3 Other exact recursions.

In Dhaene & Vandebroek (1995) the following recursion for computing the probabilities $p\,(s)$ in the individual risk model is proposed:

$$p(s) = \frac{1}{s} \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \; v_{ij}(s), \qquad s = 1, \; 2, \ldots, \tag{12}$$

with initial value given by (11) and where the coefficients $v_{ij}(s)$ are determined by

$$v_{ij}(s) = \frac{q_j}{p_j} \sum_{x=1}^{s} g_i(x) \left[x \; p(s-x) - v_{ij}(s-x) \right], \qquad s = 1, \; 2, \ldots \tag{13}$$

and $v_{ij}(s) = 0$ otherwise.

In order to derive this recursive formula, consider the probability generating function of the aggregate claims given in (4).Taking the derivative of (4) gives rise to

$$P'\,(u) = P\,(u) \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \frac{q_j G_i'\,(u)}{p_j + q_j G_i\,(u)}.$$

Since

$$G_i'\,(u) = \sum_{x=1}^{\infty} x \; g_i\,(x) \; u^{x-1} = \sum_{x=0}^{\infty} (x+1) \; g_i\,(x+1) \; u^x,$$

we have that

$$P'\,(u) = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} V_{ij}\,(u)\,, \tag{14}$$

where the auxiliary functions $V_{ij}(u)$ are defined by

$$V_{ij}(u) = \frac{q_j G_i'(u) P(u)}{p_j + q_j G_i(u)} = \sum_{x=0}^{\infty} v_{ij}(x+1)\ u^x. \tag{15}$$

The following relations hold for the derivatives of order t, with $t = 1,\ 2,\ \ldots,$

$$\begin{aligned} P^{(t)}(0) &= t!\ p_S(t), \\ V_{ij}^{(t)}(0) &= t!\ v_{ij}(t+1). \end{aligned}$$

Hence, taking the derivative of order $s-1$ of (14), taking into account Leibnitz's formula and letting $u = 0$ leads immediately to formula (12).

The recursion formula (13) for the $v_{ij}(s)$ is obtained by taking according to Leibnitz formula the derivative of order $s-1$ of the following relation, which follows from rearranging (15):

$$V_{ij}(u) = \frac{q_j}{p_j}\left(G_i'(u) P(u) - G_i(u) V_{ij}(u)\right),$$

and letting $u = 0$. This completes the proof of the recursive formula of Dhaene & Vandebroek (1995).

Alternatively, De Pril (1989) proposed the following exact recursive formula:

$$p(s) = \frac{1}{s} \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \sum_{x=1}^{s} w_{ij}(x)\ p(s-x), \qquad s = 1,\ 2,\ \ldots, \tag{16}$$

with initial value given by (11) and where the coefficients $w_{ij}(x)$ are given by

$$w_{ij}(x) = \frac{q_j}{p_j}\left(x\ g_i(x) - \sum_{k=1}^{x} g_i(k)\, w_{ij}(x-k)\right), \qquad x = 1,\ 2,\ \ldots \tag{17}$$

and $w_{ij}(x) = 0$ elsewhere.

In order to obtain this recursion, define

$$W_{ij}(u) = \frac{V_{ij}(u)}{P(u)},$$

where the $V_{ij}(u)$ are the auxiliary functions defined in (15). Hence,

$$W_{ij}(u) = \frac{q_j\ G_i'(u)}{p_j + q_j G_i(u)} = \sum_{x=0}^{\infty} w_{ij}(x+1)\ u^x.$$

Proceeding as above, we find the recursions in (16) and (17).

Dhaene & Vandebroek (1995) compare the performance of both exact procedures in (16), (17) and (12), (13). They conclude that their algorithm is more efficient than De Pril's method as to the number of multiplications to be carried out in case the portfolio

is not too heterogeneous. For very heterogeneous portfolios De Pril's method will have to be preferred.

In the following examples, we consider some special cases of the general individual risk model.

Example 1. Consider a life coverage with double indemnity provision providing a death benefit which is doubled when death is caused by an accident. The conditional claim amounts distribution is then given by

$$g_i(x) = \begin{cases} 1-\alpha & x = i \\ \alpha & x = 2i \\ 0 & \text{elsewhere} \end{cases}, \tag{18}$$

so that the aggregate claims distribution can be computed by (12) and (13) with

$$v_{ij}(s) = \frac{q_j}{p_j} \left\{(1-\alpha)\left[i\ p(s-i) - v_{ij}(s-i)\right] + \alpha\left[2\ i\ p(s-2i) - v_{ij}(s-2i)\right]\right\}. \tag{19}$$

Example 2. Consider a portfolio with all the policies having identical claim probability distributions. In this case, we can omit the indices i and j in q_j, p_j and $g_i(x)$. Furthermore, the aggregate claims have a compound binomial distribution. We obtain from (12) and (13):

$$p(s) = \frac{q}{p} \sum_{x=1}^{s} g(x) \left[\frac{(n+1)\ x}{s} - 1\right] p(s-x), \qquad s = 1,\ 2,\ \ldots, \tag{20}$$

with initial value given by

$$p(0) = p^n. \tag{21}$$

This recursion is a special case of the recursion of Panjer (1981), see also De Pril (1986).

2.4 The individual life model

The individual life model is obtained from the general individual risk model considered above by choosing

$$g_i(x) = \delta_{c_i,x}, \qquad i = 1, 2, \ldots a, \tag{22}$$

with c_i the amount-at-risk of the policies in row i and where $\delta_{a,b} = 1$ if $a = b$ and 0 otherwise.

In this case, since $g_i^{*k}(x) = \delta_{k\,c_i,x}$ the recursion in (10) reduces to

$$p(s) = \frac{1}{s} \sum_{i=1}^{a} \sum_{k=1}^{[s/c_i]} a(i,k)\ p(s - k\ c_i), \qquad s = 1,\ 2,\ \ldots, \tag{23}$$

with

$$a(i,k) = (-1)^{k+1}\, c_i \sum_{j=1}^{b} n_{ij} \left(\frac{q_j}{p_j}\right)^k, \tag{24}$$

where $[a]$ is the largest integer smaller than or equal to a, and where $\sum_{l=m}^{n}$ equals 0 if m exceeds n. This recursion was first derived in De Pril (1986).

In the individual life model, the coefficients $v_{ij}(s)$ in (13) reduce to

$$v_{ij}(s) = \frac{q_j}{p_j}\left[c_i\ p(s-c_i) - v_{ij}(s-c_i)\right], \qquad s = 1,\ 2,\ \ldots \tag{25}$$

and the recursion in (12) and (25) is due to Waldmann (1994). Waldmann obtains his results by a rearrangement of the recursive scheme of De Pril (1986) and shows that his exact recursive scheme is more efficient than the original algorithm of De Pril.

Recursions for the number of policies that lead to a strictly positive claim amount are obtained from (23) and (25) by setting all c_i equal to 1. In this case, the recursion of De Pril (1986) given in (23) reduces to the recursion of White & Greville (1959).

3 Approximate recursions

3.1 The r-th order approximations

The problem of computing the probabilities $p(s)$ numerically is not always solved satisfactory by deriving an explicit expression for the coefficients $t(x)$ as defined in (3). This will only be the case if the $t(x)$ can easily be computed, as in the individual life model. In case these coefficients are too difficult to work with, it is of interest to consider approximative methods. Following the ideas of De Pril (1988), (1989) and Dhaene & De Pril (1994), approximations $f(s)$ of the probabilities $p(s)$ are set up by replacing the exact coefficients $t(x)$ in the recursion (7) by easier computable values $h(x)$.

For a given choice of the coefficients $h(x)$ and of the starting value $f(0)$, approximations $f(s)$ of $p(s)$ are thus determined by the recursion

$$f(s) = \frac{1}{s}\sum_{x=1}^{s} x\ h(x)\ f(s-x), \qquad s = 1,\ 2,\ \ldots. \tag{26}$$

Hereafter, we introduce the coefficients $h\,(x)$ that lead to the approximative recursive methods for the individual model proposed by De Pril (1989), Hipp (1986) and Kornya (1983). The approximate algorithms perform well in case all claim probabilities q_j are sufficiently small, more precisely if $q_j < \frac{1}{2},\ j = 1, 2, \ldots, b$, which will be assumed in the sequel.

Remark that for each of the methods the approximations $f(s)$ can be negative for some values of s so that they are not necessarily probabilities anymore.

The approximations of De Pril. De Pril (1989) proposed the 'r-th order approximations' $f^{(r)}(s)$ for the 'exact probabilities' $p(s)$ by omitting in (10) all terms in $\left(\frac{q_j}{p_j}\right)^{r+1}$ and higher. This leads to the following recursion:

$$f^{(r)}(s) = \frac{1}{s}\sum_{i=1}^{a}\sum_{k=1}^{min(r,s)} A(i,k)\ \sum_{x=k}^{s} x\ g_i^{*k}(x)\ f^{(r)}(s-x) \text{ for } \ s = 1, 2, \ldots, \tag{27}$$

with starting value $f^{(r)}(0) = p(0)$. This recursion is of the form (26) with

$$\begin{aligned} h^{(r)}(0) &= t(0) = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \ln p_j, \qquad (28)\\ h^{(r)}(x) &= \sum_{i=1}^{a}\sum_{k=1}^{\min(r,x)} \frac{(-1)^{k+1}}{k}\sum_{j=1}^{b} n_{ij}\ \left(\frac{q_j}{p_j}\right)^k g_i^{*k}(x), \qquad x = 1,\ 2, \ldots. \end{aligned}$$

Observe that the r-th order approximations $f^{(r)}(s)$ in (27) are exact for the first r values.

The approximations of Kornya. The initial value $p(0) = e^{t(0)}$ with $t(0)$ defined in (8) can also be written as

$$p(0) = \exp\left(\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}\sum_{k=1}^{\infty}\frac{(-1)^k}{k}\left(\frac{q_j}{p_j}\right)^k\right). \tag{29}$$

Kornya (1983) proposes the same recursion as in (27), but with an alternative initial value $f^{(r)}(0)$ which is defined by restricting the summation over k in (29) to a summation over all k values smaller than or equal to r, with r some positive integer. Given r, the exact value $p(0)$ is thus approximated by

$$f^{(r)}(0) = \exp\left(\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}\sum_{k=1}^{r}\frac{(-1)^k}{k}\left(\frac{q_j}{p_j}\right)^k\right), \tag{30}$$

In this way, an approximative recursion of the form (26) is obtained with

$$\begin{aligned} h^{(r)}(0) &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}\sum_{k=1}^{r}\frac{(-1)^k}{k}\left(\frac{q_j}{p_j}\right)^k, \qquad (31)\\ h^{(r)}(x) &= \sum_{i=1}^{a}\sum_{k=1}^{\min(r,x)} \frac{(-1)^{k+1}}{k}\sum_{j=1}^{b} n_{ij}\ \left(\frac{q_j}{p_j}\right)^k g_i^{*k}(x), \qquad x = 1,\ 2, \ldots. \end{aligned}$$

This recursion was developed in two stages. The first stage is given in Kornya (1983), where the special case of the individual life model is considered. The generalization to

arbitrary positive claims can be found in Hipp (1986). One can easily prove that the approximations $f^{(r)}(s)$ of Kornya sum up to 1.

The approximations of Hipp. Hipp (1986) defines alternate approximate recursions for the individual model which are also of the form (26). The starting point is to write $\ln P(u)$ as

$$\begin{aligned}\ln P(u) &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \ln\left(1+q_j\left(G_i(u)-1\right)\right) = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} q_j^k \left(G_i(u)-1\right)^k \\ &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} q_j^k \sum_{l=0}^{k} \frac{(-1)^{l+1}}{(-1)^{k+1}} \binom{k}{l} \left(G_i(u)\right)^l .\end{aligned}$$

This expansion is of the form (3), with coefficients $t(x)$ given by

$$\begin{aligned} t(0) &= -\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \sum_{k=1}^{\infty} \frac{q_j^k}{k}, \qquad (32) \\ t(x) &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \sum_{k=1}^{x}\sum_{l=1}^{k} \frac{(-1)^{l+1}}{k} \binom{k}{l} q_j^k \; g_i^{*l}(x), \qquad x = 1,\ 2, \ldots, \end{aligned}$$

Let r be some positive integer. Hipp's r-th order approximations is defined by restricting the summations over k in (32) to summations over all k values smaller than or equal to r, with r some positive integer. The recursion that arises is of the form (26) with

$$\begin{aligned} h^{(r)}(0) &= -\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \sum_{k=1}^{r} \frac{q_j^k}{k}, \qquad (33) \\ h^{(r)}(x) &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \sum_{k=1}^{\min(r,x)}\sum_{l=1}^{k} \frac{(-1)^{l+1}}{k} \binom{k}{l} q_j^k \; g_i^{*l}(x), \qquad x = 1,\ 2, \ldots. \end{aligned}$$

Hipp's approximations require more computation time but have the advantage that the r-th order approximation leads to 'exact moments' up to order r, see Dhaene, Sundt & De Pril (1996).

3.2 Compound Poisson approximations

Alternatively, approximative recursive formulas for the computation of the aggregate claims distribution can be obtained by replacing the individual model by a collective model and using the recursion of Panjer (1981) to evaluate the associated compound distributions, see e.g. Bühlmann, Gagliardi, Gerber & Straub (1977), Gerber (1984), Hipp

(1985), Michel (1987), Hipp & Michel (1990) and Kuon, Radtke & Reich (1993). An overview and extensions of these results are presented in De Pril & Dhaene (1992).

A key problem to obtain a good fit between the individual and the collective model is the choice of the distribution of the number of claims and of the claims amounts. A common approximation for the individual risk model, is to replace the distribution of the claim amounts of each policy in class (i, j) by a compound Poisson distribution with parameter λ_j and severity distribution g_i. From the independence assumption, it follows that the aggregate claims S is then approximated by a compound Poisson distribution with parameter

$$\lambda = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \ \lambda_j, \tag{34}$$

and severity distribution g given by

$$g(x) = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \ \lambda_j \ g_i(x)}{\lambda}, \qquad x = 1, \ 2, \ldots, \tag{35}$$

see e.g. Gerber (1984), Hipp (1985) and De Pril & Dhaene (1992). Denoting the approximation for $p(s)$ by $f^{cP}(s)$ in this particular case, we find from Panjer's (1981) formula that the approximated probabilities can be computed from the recursion

$$f^{cP}(s) = \frac{1}{s} \sum_{x=1}^{s} x \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \ \lambda_j \ g_i(x) \ f^{cP}(s-x) \quad \text{for } s = 1, 2, \ldots, \tag{36}$$

with starting value $f^{cP}(0) = e^{-\lambda}$. This recursion is of the form (26) with

$$\begin{aligned} h\,(0) &= -\lambda, \\ h\,(x) &= \lambda g(x), \qquad x = 1, \ 2, \ldots \end{aligned}$$

The most common choice for the parameters is $\lambda_j = q_j$ which guarantees that the exact and the approximate distribution have the same expectation. This approximation, which is often referred to as *the* compound Poisson approximation is identical to the first-order approximation of Hipp (1986).

In case the Poisson parameters λ_j are chosen as

$$\lambda_j = \frac{q_j}{p_j}, \qquad j = 1, 2, \ldots, b, \tag{37}$$

the recursion in (36) corresponds with the first-order approximation of Kornya (1983).

3.3 Error bounds

As is stated in Kaas (1993), several kinds of error have to be considered when computing the aggregate claims distribution. A first type of error results from rounding the possible

claim amounts of the policies to some monetary unit, e.g. \$ 1000. A second type of error arises if the assumption of independence is violated. Computing the aggregate claims distribution of the portfolio generates a third type of error if this computation is done approximately. Hereafter, we only consider this third type of error.

For any of the approximative methods as explained in the previous subsection, let $f^{(r)}(s)$ be the r-th order approximation of $p(s)$. In Dhaene & De Pril (1994), the following result is derived which gives a quantitative measure to assess the quality of the $f^{(r)}(s)$ as approximations of the $p(s)$.
If there exists a real number $\varepsilon(r) > 0$ such that

$$\sum_{x=0}^{\infty} \left| t(x) - h^{(r)}(x) \right| \leq \varepsilon(r), \tag{38}$$

then the following error bound holds:

$$\sum_{s=0}^{\infty} \left| p(s) - f^{(r)}(s) \right| \leq e^{\varepsilon(r)} - 1. \tag{39}$$

Furthermore, Dhaene & De Pril (1994) consider approximating stop-loss premiums by means of the approximated probabilities $f^{(r)}(s)$ and derive related error bounds. Starting from the expression

$$\Pi(d) = E\left[(S-d)_+\right] = \sum_{s=0}^{d} (d-s)\ p(s) + E[S] - d, \text{ for } d = 0, 1, \ldots, \tag{40}$$

for the stop-loss premium with retention d, they propose to approximate $\Pi(d)$ by $\Pi^{(r)}(d)$ which is given by

$$\Pi^{(r)}(d) = \sum_{s=0}^{d} (d-s)\ f^{(r)}(s) + E[S] - d, \text{ for } d = 0, 1, \ldots. \tag{41}$$

Assuming that $\varepsilon(r) < \ln 2$, with $\varepsilon(r)$ defined as before, they prove the following error bound related to this approximation:

$$\left| \Pi(d) - \Pi^{(r)}(d) \right| \leq \frac{e^{\varepsilon(r)} - 1}{2 - e^{\varepsilon(r)}} \left(\Pi^{(r)}(d) + d - E[S] \right), \tag{42}$$

Error bounds related to probabilities and stop-loss premiums for De Pril's, as well as Kornya's and Hipp's approximations follow easily from (39) and (42) . For each method, the key quantities $\varepsilon(r)$ needed to compute these error bounds are deduced hereafter. In the sequel, we introduce the subscripts P, K and H to distinguish between the methods of De Pril, Kornya and Hipp.

Error bounds for De Pril's approximations. De Pril's r-th order approximations $f^{(r)}(s)$ in (27) are exact for the first r values. Hence, from (8) and (28) we find

$$\begin{aligned}
\sum_{x=0}^{\infty}\left|t(x)-h^{(r)}(x)\right| &\leq \sum_{x=r+1}^{\infty}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\sum_{k=r+1}^{x}\left|\frac{(-1)^{k+1}}{k}\left(\frac{q_j}{p_j}\right)^k g_i^{*k}(x)\right| \\
&\leq \frac{1}{r+1}\sum_{x=r+1}^{\infty}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\sum_{k=r+1}^{x}\left(\frac{q_j}{p_j}\right)^k g_i^{*k}(x) \\
&= \frac{1}{r+1}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\sum_{k=r+1}^{\infty}\left(\frac{q_j}{p_j}\right)^k\sum_{x=k}^{\infty}g_i^{*k}(x) \\
&= \frac{1}{r+1}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\frac{p_j}{p_j-q_j}\left(\frac{q_j}{p_j}\right)^{r+1}=\varepsilon_P(r). \qquad (43)
\end{aligned}$$

Error bounds for Kornya's approximations. From (8) and (31) it follows that

$$\left|t(0)-h^{(r)}(0)\right|\leq\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\left|\sum_{k=r+1}^{\infty}\frac{1}{k}\left(-\frac{q_j}{p_j}\right)^k\right|\leq\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}\,p_j}{r+1}\left(\frac{q_j}{p_j}\right)^{r+1},$$

and by proceeding as above

$$\begin{aligned}
\sum_{x=1}^{\infty}\left|t(x)-h^{(r)}(x)\right| &= \sum_{x=r+1}^{\infty}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\sum_{k=r+1}^{x}\left|\frac{(-1)^{k+1}}{k}\left(\frac{q_j}{p_j}\right)^k g_i^{*k}(x)\right| \\
&\leq \frac{1}{r+1}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\frac{p_j}{p_j-q_j}\left(\frac{q_j}{p_j}\right)^{r+1}.
\end{aligned}$$

Hence, for Kornya's approximation, we find

$$\varepsilon_K(r)=\frac{1}{r+1}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\left(p_j+\frac{p_j}{p_j-q_j}\right)\left(\frac{q_j}{p_j}\right)^{r+1}. \qquad (44)$$

Error bounds for Hipp's approximations. From (32) and (33) one has

$$\begin{aligned}
\sum_{x=0}^{\infty}\left|t(x)-h^{(r)}(x)\right| &\leq \sum_{x=0}^{\infty}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\sum_{k=r+1}^{\infty}\sum_{l=0}^{k}\left|\frac{(-1)^{l+1}}{k}\binom{k}{l}q_j^k\ g_i^{*l}(x)\right| \\
&\leq \sum_{x=0}^{\infty}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\sum_{k=r+1}^{\infty}\sum_{l=0}^{k}\frac{1}{k}\binom{k}{l}q_j^k\ g_i^{*l}(x) \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\sum_{k=r+1}^{\infty}\frac{1}{k}(2q_j)^k\leq\frac{1}{r+1}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\frac{(2q_j)^{r+1}}{p_j-q_j}=\varepsilon_H(r).
\end{aligned} \qquad (45)$$

From this result, we immediately find the following error bound for *the* compound Poisson approximation:

$$\sum_{s=0}^{\infty} \left| p(s) - f^{cP}(s) \right| \leq \exp\left(\frac{1}{2} \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \frac{(2q_j)^2}{p_j - q_j} \right) - 1. \tag{46}$$

Alternative error bounds for the general class of compound Poisson approximations defined by (36) are considered in Gerber (1984), see also De Pril & Dhaene (1992). For *the* compound Poisson approximation they deduce the following bound:

$$\sum_{s=0}^{\infty} \left| p(s) - f^{cP}(s) \right| \leq \frac{1}{2} \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \left(q_j\right)^2,$$

which is sharper than the bound given in (46). Unfortunately, the method used there to derive error bounds is specific for compound Poisson distributions and can not be generalized to the general type of approximations considered in this section.

4 Computational aspects

The efficiency of the different exact and approximate recursions presented in this paper, in terms of the number of computations required and the size of the array of the data to be kept, has been investigated partly in Dhaene & Vandebroek (1995). They conclude that if an exact recursive method is required, Dhaene-Vandebroek's recursion (12)-(13) will be the optimal choice if the portfolio is not too heterogeneous. In most practical cases however, an 'r-th order approximation' will be the best choice. From the theoretical error bounds for the r-th order approximations obtained in the previous section, but also from numerical examples, see Vandebroek & De Pril (1988), it follows that in practical situations, with values of the claim probabilities q_j smaller than $\frac{1}{2}$, the approximations for the probabilities converge very quickly to their respective exact values, so that a choice of r equal to 3, 4 or 5 will give almost exact results.
From (43), (44) and (45) it follows that

$$\varepsilon_P(r) < \varepsilon_K(r) < \varepsilon_H(r),$$

so that De Pril's approximative probabilities $f^{(r)}(s)$ give rise to smaller error bounds than the approximations proposed by Kornya and Hipp. Since the computational effort is about the same in the three cases, preference should be given to De Pril's approximations. It is important to note however that the method with the smallest error bound will not always be the best approximation as is illustrated in an example in Kornya (2004).
A comparison of the compound Poisson approximation with De Pril's, Kornya's and Hipp's methods is given in Kuon, Reich & Reimers (1987).

In this section, we will complement the results obtained in Dhaene & Vandebroek (1995) by considering computational aspects of the recursive scheme (12)-(13), of De

Pril's approximative formula (27) and also of the straightforward convolution technique (2) for the individual risk model. We will compare these three alternative methods. The comparison is based on the number of dot operations (multiplications and divisions) involved by each method, as dot operations are in general more time-consuming than summations or subtractions. For each of the three algorithms, we will count the number of dot operations to be carried out for computing $p(0)$, $p(1)$, ..., $p(s)$ for s sufficiently large. These comparisons may give an idea of which method is most efficient.

As pointed out by Kaas (1993) and Sundt & Dickson (2000), one should be aware that there are also other aspects (e.g. storage, stability etc.) that should be taken into account when choosing an algorithm. Some computational aspects of recursive formulas concerning numerical stability have been investigated in Panjer & Wang (1993), (1995).

Depending on the support of the g_i, the number of dot operations required by a particular algorithm can be significantly different. This is why we will consider two situations in the remainder of this section. First, we will consider the case that for any i, the severity function $g_i(x)$ has an unbounded support. Next, we will consider the individual life model as described in Section 2.4.

4.1 Unbounded severity distributions

In this subsection, we will assume that $g_i(x) > 0$ for all $x \geq 1$ and for all $i = 1, \ldots, a$.

4.1.1 The convolution method

Notice that in order to evaluate the distribution function of S up to value s, all the formulas involved can be stopped at s. We start by counting the dot operations required for the evaluation of the probability functions (1) of all X_{ijk}. Since for all $i = 1, \ldots, a$ and $j = 1, \ldots, b$, we only need the values $p_j, q_j g_i(1), ..., q_j g_i(s)$, this leads to exactly abs dot operations.

Considering the particularities of the convolution method, we evaluate p in two steps:

1. Evaluate each $h_{ij}^{*n_{ij}}$.
2. Evaluate $p = \overset{a}{\underset{i=1}{*}} \overset{b}{\underset{j=1}{*}} h_{ij}^{*n_{ij}}$.

In the first step we have to evaluate convolutions of higher order, which can be done in several ways, see e.g. Sundt & Dickson (2000). In the second step we have to evaluate the convolution of different distributions.

Step 1. Sundt & Dickson (2000) present the following optimal strategy to evaluate a generic f^{*u} by strictly convolution:

1. First consider the following binary representation of u :

$$u = 2^{k(u)} + \sum_{i=0}^{k(u)-1} 2^i b_{ui},$$

where $k(u)$ is a positive integer and $b_{ui} \in \{0,1\}$ for $i = 0, ..., k(u) - 1$;

2. Evaluate $f^{*2^i} = f^{*2^{i-1}} * f^{*2^{i-1}}$ for $i = 1, ..., k(u)$;

3. Finally, evaluate $f^{*u} = f^{*2^{k(u)}} * \left(\underset{\{i:b_{ui}=1\}}{*} f^{*2^i} \right)$.

Assuming that u is a power of two, i.e. $u = 2^{k(u)}$, applying the above strategy to evaluate $f^{*u}(0), ..., f^{*u}(s)$ will require $k(u)(s+2)^2/4$ dot operations if s is even, and $k(u)(s^2+4s+3)/4$ dot operations if s is odd, see Sundt & Dickson (2000).

A natural assumption is that all $n_{ij} > 1$. In order to prove that convolutions are generally more time consuming, we will consider the "best scenario", that is we assume that $n_{ij} = 2^{l_{ij}}$, where l_{ij} is a positive integer for all i, j, and we also assume that s is odd. Then all $h_{ij}^{*n_{ij}}(0), ..., h_{ij}^{*n_{ij}}(s)$ will involve

$$\frac{s^2+4s+3}{4} \sum_{i,j} l_{ij} = \frac{s^2+4s+3}{4} \log_2 \left(\prod_{i,j} n_{ij} \right) \text{ dot operations.}$$

Step 2. For distinct f and g, formula (2) involves $x+1$ dot operations. For $x = 0, ..., s$ this gives a total of

$$1 + 2 + ... + (s+1) = \frac{(s+1)(s+2)}{2} \text{ dot operations.} \tag{47}$$

Hence, evaluating $p = \underset{i=1}{\overset{a}{*}} \underset{j=1}{\overset{b}{*}} h_{ij}^{*n_{ij}}$ requires $\frac{(ab-1)(s+1)(s+2)}{2}$ dot operations.

Adding this number to the number of dot operations counted for step 1 and to abs dot operations from the evaluation of all h_{ij}, we get a total number of dot operations equal to

$$\frac{(ab-1)(s^2+3s+2)}{2} + \frac{s^2+4s+3}{4} \log_2 \left(\prod_{i,j} n_{ij} \right) + abs =$$

$$\left[\frac{1}{4} \log_2 \left(\prod_{i,j} n_{ij} \right) + \frac{ab-1}{2} \right] s^2 + \left[\log_2 \left(\prod_{i,j} n_{ij} \right) + \frac{5ab-3}{2} \right] s + \frac{3}{4} \log_2 \left(\prod_{i,j} n_{ij} \right) + ab - 1. \tag{48}$$

4.1.2 The Dhaene-Vandebroek recursion

Let us now count the number of dot operations involved in the recursion described in (12) and (13).

As can be seen from (12), calculating a single value of $p(x)$ requires $ab+1$ dot operations, so that determining $p(1), ..., p(s)$ require $(ab+1)s$ dot operations.

For a fixed s, formula (13) leads to $2s+1$ dot operations. As we have to perform this for $i = 1, ..., a,\ j = 1, ..., b$, i.e. ab times, we find a total of $ab(2s+1)$ dot operations. We also need to evaluate $v_{ij}(1), ..., v_{ij}(s)$, which requires $\sum_{i=1}^{s} ab(2i+1) = ab[s(s+1)+s]$ dot operations. Notice that we also have to evaluate q_j/p_j for each j, which totally gives b dot operations.

Combining these figures, we can conclude that the Dhaene-Vandebroek algorithm involves

$$s(ab+1) + ab[s(s+1)+s] + b = abs^2 + (3ab+1)s + b \tag{49}$$

dot operations.

4.1.3 De Pril's r-th order approximation

Finally, we consider the number of dot operations involved in De Pril's r-th order approximation (27). We assume that r is sufficiently small such that $r < s$.

Starting with the convolutions $g_i^{*k}(x)$ in (27), we see that we need all the convolutions from $g_i^{*2}(2)$ till $g_i^{*r}(s), i = 1, ..., a$. Since $g_i(x)$ is defined for $x \geq 1$, we notice that

$$g_i^{*k}(y) = \sum_{x=1}^{y-k+1} g_i^{*(k-1)}(y-x) g_i(x), \ \ k = 2, ..., y, \tag{50}$$

because we must have that $y - x \geq k - 1$. Hence determining $g_i^{*k}(y)$ requires $y-k+1$ dot operations. Summing over y and k, we obtain a total number of dot operations given by

$$\begin{aligned}
\sum_{k=2}^{r}\sum_{y=k}^{s}(y-k+1) &= \sum_{k=2}^{r}\sum_{z=1}^{s-k+1} z = \sum_{k=2}^{r}\frac{(s-k+1)(s-k+2)}{2} = \sum_{j=1}^{r-1}\frac{(s-j)(s-j+1)}{2} \\
&= \frac{1}{2}\left[\left(s^2+s\right)(r-1) + \frac{(r-1)r(2r-1)}{6} - (2s+1)\frac{(r-1)r}{2}\right] \\
&= \frac{r-1}{2}\left[s^2 - (r-1)s + \frac{(r-2)r}{3}\right].
\end{aligned}$$

Considering now $i = 1, ..., a$, we obtain a total of $\dfrac{a(r-1)}{2}\left[s^2 - (r-1)s + \dfrac{(r-2)r}{3}\right]$ dot operations. Since we need high order convolutions for the same distribution function, we could apply the more efficient technique presented in Section 8 of Sundt & Dickson (2000). However, since r is small, the difference will not be that important.

Since the evaluation of the $A(i,k)$ does not involve s, we neglect it for the moment.

Let us now consider the calculation of the $f^{(r)}$ as given by (27). In order to reduce the number of dot operations, we rewrite (27) as

$$f^{(r)}(y) = \frac{1}{y}\sum_{x=1}^{y} f^{(r)}(y-x)x \sum_{k=1}^{\min(r,x,y)} \sum_{i=1}^{a} A(i,k)\; g_i^{*k}(x) = \frac{1}{y}\sum_{x=1}^{y} \varphi_{f^{(r)}}(x)\, f^{(r)}(y-x)\,, \quad (51)$$

where

$$\varphi_{f^{(r)}}(x) = x \sum_{i=1}^{a} \sum_{k=1}^{\min(r,x)} A(i,k)\, g_i^{*k}(x)\,.$$

Since we need $f^{(r)}$ up to value s, we see that one should first evaluate all $\varphi_{f^{(r)}}(1), ..., \varphi_{f^{(r)}}(s)$ separately. This implies two cases:

1. In case $x < r$, we have that $\varphi_{f^{(r)}}(x)$ needs a single dot operation from the multiplication with x and ax operations for the products in the double sum.

2. In case $x \geq r$, we have that $\varphi_{f^{(r)}}(x)$ needs one dot operation from the multiplication with x and ar operations for the products in the double sum.

Calculating $\varphi_{f^{(r)}}(1), ..., \varphi_{f^{(r)}}(s)$ will then require

$$\begin{aligned}\sum_{x=1}^{r-1}(ax+1) + \sum_{x=r}^{s}(ar+1) &= a\frac{(r-1)\,r}{2} + r - 1 + (ar+1)(s-r+1) \\ &= (ar+1)\,s - \frac{ar}{2}(r-1)\end{aligned}$$

dot operations. We also see from (51) that a single $f^{(r)}(y)$ involves $y+1$ dot operations, so that $f^{(r)}(1), ..., f^{(r)}(s)$ will add $\sum_{y=1}^{s}(y+1) = \frac{s\,(s+1)}{2} + s$ to the number above. This gives rise to a total of

$$\frac{1}{2}s^2 + \left(ar + \frac{5}{2}\right)s - \frac{ar}{2}(r-1)$$

dot operations. Finally, we have to add the dot operations for the evaluation of the convolutions g_i^{*k}. Hence, for the r-th order De Pril approximation we need a total amount of dot operations given by

$$\begin{aligned}&\frac{a\,(r-1)}{2}\left[s^2 - (r-1)\,s + \frac{(r-2)\,r}{3}\right] + \frac{1}{2}s^2 + \left(ar + \frac{5}{2}\right)s - \frac{ar}{2}(r-1) \\ &= \frac{a\,(r-1)+1}{2}s^2 + \left[\left(r - \frac{(r-1)^2}{2}\right)a + \frac{5}{2}\right]s + \frac{ar\,(r-1)\,(r-5)}{6}. \qquad (52)\end{aligned}$$

Here, we should add about abr dot operations for the evaluation of the $A(i,k), i = 1, ..., a,\ k = 1, ..., r$ and b for evaluating q_j/p_j for each j.

4.1.4 Conclusions for the unbounded case

The results in the previous subsections indicate that the evaluation of the distribution function of S up to value s requires a number of dot operations depending on s^2 for all three algorithms. We will now compare the three methods, mainly based on comparing the respective coefficients of s^2.

Notice that for the convolution method, we will consider the number of dot operations (48) related with the 'best scenario' as explained earlier.

Since the algorithm of De Pril is in fact an approximative one, whereas the two others are exact, we will indicate for what values of r De Pril's approximation method requires less dot operations than the exact methods.

1. **Convolutions versus Dhaene-Vandebroek**. Taking into consideration (49) and (48), it is easy to see that the Dhaene-Vandebroek algorithm requires less dot operations when $\prod_{i,j} n_{ij} > 4^{ab+1}$. This inequality indicates that for a fixed number of classes, i.e. for a fixed value of a and b, increasing the number of members in the different classes will lead to a better performance of the Dhaene-Vandebroek algorithm, relative to the convolution method. This implies that a sufficient level of homogeneity in the portfolio, in the sense that n is large compared to ab, improves the performance of the Dhaene-Vandebroek algorithm. This conclusion could be expected from (12) and (13) which indicate that the Dhaene-Vandebroek method benefits from having several members in one cell, while this is not the case for the convolution method.

2. **Convolutions versus De Pril**. From (48) and (52) we can conclude that De Pril performs better compared to the straightforward convolution method provided $r \leq b + 1 - \frac{2}{a} + \frac{1}{2a} \log_2 \left(\prod_{i,j} n_{ij} \right)$. For r sufficiently small, this condition will be satisfied.

3. **Dhaene-Vandebroek versus De Pril**. Comparing (49) and (52), we see that De Pril performs better when $r \leq 2b + 1 - a^{-1}$, which will be satisfied when r is sufficiently small. This supports the conclusion of Dhaene & Vandebroek (1995).

We can conclude that under reasonable assumptions (i.e. for n large enough), when evaluating the distribution function, the convolution method will require a larger number of dot operations than the other two methods.

4.2 The individual life model

Next, we consider the individual life model (22) presented in Section 2.2. Let $n_{i+} = \sum_{j=1}^{b} n_{ij}$ be the number of policies with severity distribution g_i. Since g_i is only strictly positive

for a single value c_i, $i = 1, ..., a$, we have that s can only take values between 0 and c_+, where $c_+ = \sum_{i=1}^{a} c_i n_{i+}$.

Depending on the particular values of the c_i, it is possible for the aggregate claims probabilities to be zero for some values between 0 and c_+. This may give rise to a reduction of the required number of dot operations. However, we will not take into account this reduction. Hence, we will count the number of operations as if p was non-zero for all values in $\{0, , 1, 2, ..., c_+\}$. In particular, this will be the case when all c_i equal 1. Further, we will also restrict to the case when $s = c_+$. Hence, we will count the dot operations needed to evaluate the entire aggregate claims distribution.

4.2.1 The convolution method

In order to determine p, we will use the optimal method as described in Section 4.1.1. Hence, consider a generic function h defined on $\{0, c, ..., mc\}$. In order to evaluate $h^{*2}(0)$, $h^{*2}(c)$, ..., $h^{*2}(2mc)$, we need all the products $h(xc)\ h(yc)$ for $0 \leq x \leq y \leq m$. This gives rise to a number of dot operations equal to

$$\sum_{y=0}^{m} (y+1) = \frac{(m+1)(m+2)}{2}.$$

Let us now assume that f is defined on the support $\{0, c\}$. In order to evaluate $f^{*u}(0), f^{*u}(c), ..., f^{*u}(cu)$, where $u = 2^l$, and using the relation $f^{*2^i} = f^{*2^{i-1}} * f^{*2^{i-1}}$, $i = 1, ..., l$, as in Section 4.1.1, we find that the number of required dot operations is given by

$$\begin{aligned} \sum_{i=1}^{l} \frac{(2^{i-1}+1)(2^{i-1}+2)}{2} &= \sum_{i=1}^{l} \frac{4^{i-1} + 3 \cdot 2^{i-1}}{2} + l = \frac{4^l - 1}{6} + \frac{3\left(2^l - 1\right)}{2} + l \\ &= \frac{u^2}{6} + \frac{3u}{2} + \log_2 u - \frac{5}{3}. \end{aligned}$$

Applying this result to all $h_{ij}^{*n_{ij}}(0)$, $h_{ij}^{*n_{ij}}(c_i)$, ..., $h_{ij}^{*n_{ij}}(c_i n_{ij})$, $i = 1, ..., a$, $j = 1, ..., b$, we obtain a minimum number of dot operations given by

$$\frac{1}{6} \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}^2 + \frac{3}{2} n + \log_2 \left(\prod_{i,j} n_{ij} \right) - \frac{5}{3} ab.$$

We still have to count the required dot operations for determining $p = \underset{i=1}{\overset{a}{*}} \underset{j=1}{\overset{b}{*}} h_{ij}^{*n_{ij}}$. For two generic probability functions f and g, where f is defined on $\{0, 1, ..., n_1\}$ and g on $\{0, c, ..., cn_2\}$, we have that the evaluation of $(f * g)(0)$, ..., $(f * g)(n_1 + cn_2)$ will involve all the products $f(x)\ g(cy)$, $x = 0, ..., n_1$, $y = 0, ..., n_2$. This will require $(n_1 + 1)(n_2 + 1)$ dot operations. Applying this reasoning to p, we find that

$$\sum_{i=1}^{a} \sum_{j=1}^{b} (n_{ij} + 1) \left(c_1 n_{1+} + ... + c_{i-1} n_{i-1+} + c_i \sum_{k=1}^{j-1} n_{ik} + 1 \right) - (n_{11} + 1)$$

dot operations are required. Adding this number to the number of operations obtained above, the total number of dot operations is found to be

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{6}+\frac{3n}{2}+\log_2\left(\prod_{i,j}n_{ij}\right)-\frac{5ab}{3}+\sum_{i=1}^{a}\sum_{j=1}^{b}(n_{ij}+1)\left(\sum_{k=1}^{i-1}c_k n_{k+}+c_i\sum_{k=1}^{j-1}n_{ik}+1\right)-n_{11}-1. \tag{53}$$

The value (53) is minimal when $a = 1$ and $c_1 = 1$. We will restrict to this case in the remainder of this subsection. In order to simplify the notations, we denote n_{1j} by n_j for any $j = 1, ..., b$. Then, (53) reduces to

$$\begin{aligned}&\frac{1}{6}\sum_{j=1}^{b}n_j^2+\sum_{1\le k<i\le b}n_i n_k+\sum_{i=2}^{b}\sum_{k=1}^{i-1}n_k+\frac{5}{2}n-n_1+\log_2\left(\prod_{i=1}^{b}n_i\right)-\frac{2}{3}b-1 \qquad (54)\\ =\ &\frac{1}{6}n^2+\frac{5}{2}n+\frac{2}{3}\sum_{1\le k<i\le b}n_i n_k+\sum_{i=2}^{b}\sum_{k=1}^{i-1}n_k+\log_2\left(\prod_{i=1}^{b}n_i\right)-n_1-\frac{2}{3}b-1.\end{aligned}$$

4.2.2 The Dhaene-Vandebroek recursion

We start counting the number of dot operations by considering formula (12), which has to be performed for $s = 1, 2, ..., c_+$. This leads to $(1 + ab)\, c_+$ dot operations.

For fixed values of i and j, formula (25) has to be performed for $s = c_i$, $c_i + 1$, ..., c_+. This leads to $2\,(c_+ - c_i + 1)$ dot operations. For all $i = 1, ..., a$, $j = 1, ..., b$, we obtain a total of

$$2b\sum_{i=1}^{a}(c_+ - c_i + 1) = 2abc_+ - 2b\sum_{i=1}^{a}c_i + 2ab$$

dot operations. The total number of dot operations required for the Dhaene-Vandebroek recursion for the individual life model is then given by

$$(1 + 3ab)\, c_+ - 2b\sum_{i=1}^{a}c_i + 2ab + b, \tag{55}$$

where the evaluation of q_j/p_j for each j, leads to additional number of b dot operations.

In the particular case when $a = 1$ and $c_1 = 1$, we notice that $c_+ = n_{1+} = n$, and (55) reduces to

$$(1 + 3b)\, n + b. \tag{56}$$

4.2.3 De Pril's r-th order approximation

For the individual risk model, the r-th order recursion (23) reduces to

$$p(s) = \frac{1}{s}\sum_{i=1}^{a}\sum_{k=1}^{\min\{r,[s/c_i]\}}a(i,k)\; p(s - kc_i), \qquad s = 1,\ 2,\ \ldots, c_+, \tag{57}$$

with $a\,(i,k)$ given in (24).

Starting with the evaluation of q_j/p_j for each j and of all $a(i,k)$, a total number of about $b + ab + abr$ dot operations will be required. For a fixed i, the number of dot operations involved in the second sum of (57) depends on the value of s. In particular, we find:

1. For $1 \leq s < c_i$, the second sum is $\sum_{k=1}^{0} = 0$.
2. For each s, $c_i \leq s < 2c_i$, the second sum is $\sum_{k=1}^{1}$, hence one dot operation.
3. For each s, $2c_i \leq s < 3c_i$, the second sum is $\sum_{k=1}^{2}$, hence 2 dot operations.

   ...
4. For each s, $rc_i \leq s \leq c_+$, the second sum is $\sum_{k=1}^{r}$, hence r dot operations.

Then, for a fixed i, the second sum in (57) requires

$$c_i + 2c_i + ... + (r-1)\,c_i + r\,(c_+ - rc_i + 1) = \frac{r\,(r-1)}{2}c_i + rc_+ + r\,(1 - rc_i)$$

dot operations. Summing for all $i = 1, ..., a$, this gives rise to

$$arc_+ + \left[\frac{r\,(r-1)}{2} - r^2\right]\sum_{i=1}^{a} c_i + ar$$

dot operations. In addition, we have c_+ more dot operation from the multiplications with $1/s$.

We can conclude that calculating all $p\,(s)$, $s = 1,\, 2,\, \ldots, c_+$, will involve a number of dot operations equal to

$$(1 + ar)\,c_+ - \frac{r\,(r+1)}{2}\sum_{i=1}^{a} c_i + ar\,(1+b) + (a+1)\,b. \tag{58}$$

In the special case that $a = 1$ and $c_1 = 1$, this relation reduces to

$$(1 + r)\,n - \frac{r\,(r+1)}{2} + r\,(1+b) + 2b. \tag{59}$$

4.2.4 Conclusions for the individual life model

From (53), (55) and (58) we see that for n sufficiently large, the number of required dot operations in the convolution method is the largest since (53) depends on n^2, while the other relations only depend on $c_+ = \sum_{i=1}^{a} c_i n_{i+}$. We can can conclude that when the number of policies involved is large enough, the convolution method will require more dot operations than any of the other methods.

Comparing (55) and (58), we have that the number of dot operations required by the Dhaene-Vandebroek recursion for the individual life model is larger than the number required by De Pril's r-th order method if

$$(1+3ab)\,c_+ - 2b\sum_{i=1}^{a} c_i + 2ab + b > (1+ar)\,c_+ - \frac{r\,(r+1)}{2}\sum_{i=1}^{a} c_i + ar\,(1+b) + (a+1)\,b,$$

which is equivalent to

$$(3b-r)\,ac_+ - \left[2b - \frac{r\,(r+1)}{2}\right]\sum_{i=1}^{a} c_i > ar\,(1+b) - ab. \tag{60}$$

Rewriting c_+ as

$$c_+ = \sum_{i=1}^{a} c_i n_{i+} = \sum_{i=1}^{a} c_i + \sum_{i=1}^{a} c_i\,(n_{i+}-1)\,,$$

the condition (60) becomes

$$(3b-r)\,a\sum_{i=1}^{a} c_i + (3b-r)\,a\sum_{i=1}^{a} c_i\,(n_{i+}-1) - \left[2b - \frac{r\,(r+1)}{2}\right]\sum_{i=1}^{a} c_i - a\,[r+b\,(r-1)] > 0,$$

or, equivalently,

$$\left[(3b-r)\,a - 2b + \frac{r\,(r+1)}{2}\right]\sum_{i=1}^{a} c_i + \left[(3b-r)\sum_{i=1}^{a} c_i\,(n_{i+}-1) - r - b\,(r-1)\right]a > 0. \tag{61}$$

A sufficient condition for this inequality to hold is that each term in (61) is positive. The first term is positive when

$$(3b-r)\,a - 2b + \frac{r\,(r+1)}{2} > 0 \Leftrightarrow 2b\,(3a-2) > r\,(2a-r-1)\,.$$

As $2a-r-1 < 3a-2$ always holds, a sufficient condition is that $r < 2b$.
Further, since $r < 2b$ implies $r < 3b$, we have that the second term of (61) is positive for n sufficiently large.

To conclude, the condition $r < 2b$ implies that the r-th order approximation method of De Pril outperforms the Dhaene-Vandebroek algorithm.

We will now have a closer look at the particular case $a = 1, c_1 = 1$. The following comparison is approximative.

1. **Convolutions versus Dhaene-Vandebroek**. In order to compare both methods, in (54) and (56), we only retain the terms in n and n^2.We find that Dhaene-Vandebroek is to be preferred in case $\frac{1}{6}n^2 + \frac{5}{2}n > (1+3b)\,n$, or equivalently,

$$18b < n + 9.$$

This condition will be satisfied for large n.

2. **Convolutions versus De Pril**. Only retaining the terms in n and n^2, from (54) and (59), we find that De Pril's r-th order method outperforms the direct convolution method in case $\frac{1}{6}n^2 + \frac{5}{2}n > (1+r)\,n$ or

$$r < \frac{n+9}{6}.$$

This condition will also be satisfied for large n.

3. **Dhaene-Vandebroek versus De Pril**. From (56) and (59), we see that De Pril requires less dot operations than Dhaene-Vandebroek if

$$(1+r)\,n - \frac{r\,(r+1)}{2} + r\,(1+b) + 2b < (1+3b)\,n + b,$$

or, equivalently,

$$(r+1)\,b - \frac{r\,(r-1)}{2} < (3b-r)\,n.$$

A sufficient condition for this to hold is that $r < 3b$.

5 Final remarks

In this paper, we presented and compared several exact and approximative recursive methods that have been presented in the actuarial literature for calculating the aggregate claims distribution in the individual risk model. We complemented the work of Dhaene & Vandebroek (1995) by comparing the number of dot operations involved in each method.

The general conclusion is that if an exact procedure is required, the Dhaene-Vandebroek algorithm will have to be preferred over straightforward convolution in case the portfolio is not too heterogenous, in the sense that the number of policies n is large compared to the number of cells ab. The reason is that the former algorithm benefits from having different policies in the same class $(i,\, j)$ while this is not the case for the convolution method.

When an approximative method is allowed, De Pril's r-th order method will have to be preferred, provided r is sufficiently small. Notice, that from a computational point of view De Pril's first order method is equivalent with a compound Poisson approximation for the individual risk model.

The one-dimensional recursions and error bounds presented in this paper can be generalized to the multivariate case arising from introducing multivariate severities, see Sundt (2000b).

A central assumption we made in this paper is that mutual independence of the claims related to the different policies in the portfolio. This assumption is relaxed in Ribas, Marín-Solano & Alegre (2003), where bivariate dependencies are taken into account.

References

[1] Bowers N.L.; Gerber, H.U.; Hickman, J.C.; Jones, D.A.; Nesbitt, C.J. (1997). *Actuarial Mathematics*, second edition, The Society of Actuaries, Schaumburg, Illinois.

[2] Bühlmann, H.; Gagliardi, B.; Gerber, H.U.; Straub, E. (1977). Some inequalities for stop-loss premiums. *ASTIN Bulletin* 9, 75-83.

[3] De Pril, N. (1985). Recursions for convolutions of arithmetic distributions, *ASTIN Bulletin* 15, 135-139.

[4] De Pril, N. (1986). On the exact computation of the aggregate claims distribution in the individual life model, *ASTIN Bulletin* 16, 109-112.

[5] De Pril, N. (1988). Improved approximations for the aggregate claims distribution of a life insurance portfolio, *Scandinavian Actuarial Journal*, 61-68.

[6] De Pril, N. (1989). The aggregate claims distribution in the individual model with arbitrary positive claims, *ASTIN Bulletin* 19, 9-24.

[7] De Pril, N. ; Dhaene, J. (1992). Error bounds for compound Poisson approximations of the individual risk model, *ASTIN Bulletin* 19(1), 135-148.

[8] Dhaene, J.; De Pril, N. (1994). On a class of approximative computation methods in the individual risk model, *Insurance: Mathematics & Economics* 14, 181-196.

[9] Dhaene, J.; Sundt, B. (1997). On error bounds for approximations to aggregate claims distributions, *ASTIN Bulletin* 27, 243-262.

[10] Dhaene, J.; Sundt, B. (1998). On approximating distributions by approximating their De Pril transforms, *Scandinavian Actuarial Journal* 1998, 1-23.

[11] Dhaene, J.; Sundt, B.; De Pril, N. (1996). Some moment relations for the Hipp approximation *ASTIN Bulletin* 26, 117-121.

[12] Dhaene, J.; Vandebroek, M. (1995). Recursions for the individual model, *Insurance: Mathematics & Economics* 16, 31-38.

[13] Dhaene, J.; Willmot, G.; Sundt, B. (1999). Recursions for distribution functions and stop-loss transforms, *Scandinavian Actuarial Journal* 1999, 52-65.

[14] Gerber, H.U. (1984). Error bounds for the compound Poisson approximation, *Insurance: Mathematics & Economics* 3, 191-194.

[15] Hipp, C. (1985). Approximation of aggregate claims distributions by compound Poisson distributions, *Insurance: Mathematics & Economics* 4, 227-232. Correction note: *Insurance: Mathematics & Economics* 6, 165.

[16] Hipp, C. (1986). Improved approximations for the aggregate claims distribution in the individual model, *ASTIN Bulletin* 16(2), 89-100.

[17] Hipp, C.; Michel, R. (1990). *Risicotheorie: Stochastische Modelle und Statistische Methode*n. Schriftenreihe Angewandte Versicherungsmathematik, Heft 24, Deutsche Gesellschaft für Versicherungsmathematik, distributed by Verlag Versicherungswirtschaft e.V., Karlsruhe.

[18] Kaas, R. (1993). How to (and how not to) compute stop-loss premiums in practice, *Insurance: Mathematics & Economics* 13, 241-254.

[19] Kaas, R.; Goovaerts, M.J.; Dhaene, J.; Denuit, M. (2001). *Modern Actuarial Risk Theory*, Kluwer Academic Publishers, Boston.

[20] Kornya, P.S. (1983) Distribution of aggregate claims in the individual risk model, *Transactions of the Society of Actuaries* 35, 823-836, Discussion: 837-858.

[21] Kornya, P.S. (2004). Approximating the individual risk model: a counterexample, unpublished notes.

[22] Kuon, S.; Reich, A.; Reimers, L. (1987) Panjer vs. Kornya vs. De Pril: A comparison from a practical point of view, *ASTIN Bulletin* 17(2), 183-191.

[23] Kuon, S.; Radtke, M.; Reich, A. (1993). An appropriate way to switch from the individual model to the collective one. *ASTIN Bulletin* 23, 23-54.

[24] Michel, R. (1987). An improved error bound for the compound Poisson approximation. *ASTIN Bulletin* 17, 165-169.

[25] Panjer, H.H. (1981) Recursive evaluation of a family of compound distributions, *ASTIN Bulletin* 12, 22-26.

[26] Panjer, H.; Wang, S. (1993). On the stability of recursive formulas, *ASTIN Bulletin* 23, 227-258.

[27] Panjer, H.; Wang, S. (1995). Computational aspects of Sundt's generalized class, *ASTIN Bulletin* 25, 5-17.

[28] Ribas, C.; Marín-Solano, J.; Alegre A. (2003). On the computation of the aggregate claims distribution in the individual life model with bivariate dependencies, *Insurance: Mathematics & Economics* 32, 201-216.

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

[30] Sundt, B. (1995). On some properties of De Pril transforms of counting distributions, *ASTIN Bulletin* 25, 19-31.

[31] Sundt, B. (1998). A generalization of the De Pril transform, *Scandinavian Actuarial Journal* 1998, 41-48.

[32] Sundt, B. (2000a). The multivariate De Pril transform, *Insurance: Mathematics & Economics* 27, 123-136.

[33] Sundt, B. (2000b). On error bounds for approximations to multivariate distributions, *Insurance: Mathematics & Economics* 27, 137-144.

[34] Sundt, B. (2002). Recursive evaluation of aggregate claims distributions, *Insurance: Mathematics & Economics* 30, 297-322.

[35] Sundt, B.; Dhaene, J.; De Pril, N. (1998). Some results on moments and cumulants, *Scandinavian Actuarial Journal* 1998, 24-40.

[36] Sundt, B.; Dickson, D.C.M. (2000). Comparison of methods for evaluation of the n-fold convolution of an arithmetic distribution. *Bulletin of the Association of Swiss Actuaries 2000(2)*, 129-140.

[37] Sundt, B.; Jewell, W.S. (1981). Further results on recursive evaluation of compound distributions, *Astin Bulletin* 13, 27-39.

[38] Vandebroek, M.; De Pril, N. (1988) Recursions for the distribution of a life portfolio: a numerical comparison, *Bulletin of the Royal Association of Belgian Actuaries* 82, 21-35.

[39] Waldmann, K.-H. (1994) On the exact calculation of the aggregate claims distribution in the individual life model, *ASTIN Bulletin* 24, 89-96.

[40] White, R.P.; Greville, T.N.E. (1959). On computing the probability that exactly k out of n independent events will occur, *Transactions of the Society of Actuaries* 11, 88-99. Discussion 96-99.