Some comments on the individual risk model and multivariate extension

J.F. Walhin[*•] and J. Paris[*]

[*] *Université Catholique de Louvain*

[•] *Secura Belgian Re*

Abstract

In this paper we will first give some comments on the individual risk model concentrating on the application and the power of the De Pril's approximation.
We extend the exact result of Dhaene and Vandebroek (1995) on the individual risk model to a multivariate setting.
We derive an approximate algorithm which is the extension of the De Pril's algorithm obtained in Sundt (1999) and in Dickson and Waters (1999). Approximate formulae are given and compared to exact ones.
Excess of loss reinsurance with reinstatements is used as an application.

Keywords

Individual risk model ; exact formula ; approximate formula ; recursive algorithm ; number of arithmetic operations ; probability generating function ; reinstatements.

1. Introduction

The individual risk model received great attention during the 80's. Kornya (1983) gave an approximate formula to solve the individual life model while De Pril (1986) gave an exact formula as well as an approximate formula based on the exact one (De Pril (1988)).
Those formulae have been extended to the individual risk model by De Pril (1989).
Recently, two very interesting papers appeared which have motivated the present paper.
Dhaene and De Pril (1994) have given a general methodology to compare the different approximations for the individual risk model.
Dhaene and Vandebroek (1995) have given a new exact recursion for the individual risk model improving the formula of De Pril (1989). They prove that this new formula performs better than the approximate formulae in case of a very heterogeneous portfolio.
The model and notations are introduced in section 2.
The De Pril's approximation is studied in the case where claims with no cost may occur (section 3). The formulation of the De Pril's approximation as a double stage algorithm proves

that the De Pril's approximation is in fact by far better than initially stated in Dhaene and Vandebroek (1995).
The rest of the paper is devoted to the extension of the methodologies in a bivariate setting. Recently, Dickson and Waters (1999) and Sundt (1999) have extended the De Pril's approximation to the multivariate setting. The main contribution of this paper is to extend the result of Dhaene and Vandebroek (1995) to a multivariate setting and to compare the different results. We will see that the error bound for the approximative De Pril's formula remains the same and we will study the time of calculation. Section 5 presents the bivariate model. Section 6 gives an approximate formula based on De Pril's approximate formula. Section 7 gives an exact formula based on Dhaene and Vandebroek's exact formula. The formulae are compared as to the time calculation in section 8. Section 9 shows that the collective risk model is also applicable in this bivariate setting. Section 10 is devoted to an application : excess of loss reinsurance with reinstatements. The extension of De Pril's formula (1985) for the convolution of arithmetic distributions in a bivariate setting is given in section 11. Section 12 gives the conclusion.

2. Model description and notations

We assume a portfolio of n independent policies. Each policy may produce one claim at the most during the period considered (typically one year). This is obviously the case in life insurance and is also quite realistic for e.g. fire insurance.
The policies are classified into $a \times b$ classes in the following way : the policies of row j have a probability q_j $(j = 1, \cdots, b)$ of producing a claim and the conditional distribution of the claim for a policy in column i given that a claim has occurred is $g_i(x)\,(i = 1, \cdots, a)$.
Let n_{ij} be the number of policies with characteristics i and j and n be the total number of policies.

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

Table 1: The individual risk model : homogeneous portfolio notation

We have

$$\sum_{j=1}^{b} \sum_{i=1}^{a} n_{ij} = n$$

It is assumed that $0 < q_j < 1$ and that the claim amounts are distributed on $x = 1, 2, \cdots, m_i$, i.e. with a maximal claim amount m_i.

In the next section, we will study what happens if one authorizes a positive mass in 0, i.e. $g_i(0) > 0$.
$M = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} m_i$ is the maximal amount of aggregate claims.
Let $p(s), s = 0, \cdots, M$ be the probability function of the aggregate claims.
In order to find this distribution, we will use the probability generating function :

$$P(u) = \sum_{s=0}^{M} p(s) u^s = \prod_{i=1}^{a} \prod_{j=1}^{b} (p_j + q_j G_i(u))^{n_{ij}}$$

where $G_i(u) = \sum_{x=1}^{m_i} g_i(x) u^x$.

Different techniques have been described in order to evaluate $p(s)$. De Pril (1989) has proposed an approximate algorithm valid if $q_j < \frac{1}{2} \quad \forall j$: one chooses the degree of approximation r and evaluates

$$\begin{aligned}
p^{(r)}(0) &= \prod_{i=1}^{a} \prod_{j=1}^{b} (1 - q_j)^{n_{ij}} \\
sp^{(r)}(s) &= \sum_{i=1}^{a} \sum_{k=1}^{\min(r,s)} A(i,k) \sum_{x=k}^{\min(s,km_i)} x g_i^{*k}(x) p^{(r)}(s-x) \quad 1 \leq s \leq M \\
A(i,k) &= \frac{(-1)^{k+1}}{k} \sum_{j=1}^{b} n_{ij} \left(\frac{q_j}{p_j} \right)^k
\end{aligned}$$

The De Pril's approximation can also be described as a double stage algorithm :

$$\begin{aligned}
t^{(r)}(0) &= \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \ln p_j \\
t^{(r)}(x) &= \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \sum_{k=1}^{\min(x,r)} \frac{(-1)^{k+1}}{k} \left[\frac{q_j}{p_j} \right]^k g_i^{*k}(x) \quad 1 \leq x \leq r \max_i m_i \\
p^{(r)}(0) &= e^{t^{(r)}(0)} \\
sp^{(r)}(s) &= \sum_{x=1}^{s} x t^{(r)}(x) p^{(r)}(s-x) \quad 1 \leq s \leq M
\end{aligned}$$

The approximation is better as r increases.
Moreover, Dhaene and De Pril (1994) have given an upper bound for the error made when approximating the true distribution of $p(s)$.

$$\sum_{s=0}^{\infty} |p(s) - p^{(r)}(s)| \leq e^{\epsilon} - 1$$

For the particular case of the De Pril's approximation, we have

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

Dhaene and Vandebroek (1995) have given an exact algorithm for the evaluation of the $p(s)$ with the following steps :

$$
\begin{aligned}
p(0) &= \prod_{i=1}^{a}\prod_{j=1}^{b}(1-q_j)^{n_{ij}} \\
sp(s) &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}v_{ij}(s) \quad 1 \le s \le M \\
v_{ij}(s) &= \frac{q_j}{p_j}\sum_{x=1}^{m_i} g_i(x)\left(xp(s-x) - v_{ij}(s-x)\right) \quad s \ge 1 \\
& 0 \quad \text{otherwise}
\end{aligned}
$$

In Dhaene and Vandebroek (1995) they compare the number of multiplications required in order to perform the exact algorithm and the approximate De Pril's one stage algorithm. They find :

$$
\begin{aligned}
\text{exact :} &\qquad 2b(a + \sum_{i=1}^{a} m_i)s \\
\text{approximate :} &\qquad \left(3ar + r(r+1)\sum_{i=1}^{a}(m_i - 1)\right)s
\end{aligned}
$$

with some initial work neglected because not depending on s for the approximate formula. They conclude that the approximate formula is less time consuming than the exact one if

$$b > \frac{r(r+1)}{2}$$

which is satisfied unless the portfolio is very heterogeneous.

3. How to use the De Pril's approximation if $g_i(0) > 0$

In this section, we allow a positive mass in 0 for the claim amount distribution. This situation may occur when a claim reported to the company does not result in any future payment e.g. because the claim amount remains under the deductible.
In this case, it is not difficult to see that

$$\ln P(u) = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \ln(p_j+q_jG_i(0))+\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \left[\frac{q_j}{p_j+q_jG_i(0)}\right]^k (G_i(u)-G_i(0))^k$$

In the double step setting described in De Pril and Dhaene (1992), we find :

$$
\begin{aligned}
t(0) &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \ln(p_j+q_jg_i(0)) \\
t(x) &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \left[\frac{q_j}{p_j+q_jg_i(0)}\right]^k (g_i(x)-g_i(0))^{*k} \quad x \geq 1 \\
p(0) &= e^{t(0)} \\
sp(s) &= \sum_{x=1}^{s} xt(x)p(s-x) \quad 1 \leq s \leq M
\end{aligned}
\tag{2}
$$

In this case, the De Pril's approximation at level r is given by

$$t^{(r)}(x) = \sum_{k=1}^{r} \frac{(-1)^{k+1}}{k} \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \left[\frac{q_j}{p_j+q_jg_i(0)}\right]^k (g_i(x)-g_i(0))^{*k} \quad 1 \leq x \leq r \max_i m_i$$

In this setting we immediately find

$$\sum_{x=0}^{\infty} |t(x)-t^{(r)}(x)| = \epsilon \leq \frac{1}{r+1} \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \frac{1-q_j(1-g_i(0))}{1-2q_j+q_jg_i(0)} \left[\frac{q_j}{1-q_j(1-g_i(0))}\right]^{r+1} \tag{3}$$

which is quite similar to the bound (1).
Now, if one authorizes $g_i(0) > 0$, it is not difficult to transform the present model into the previous one where $g_i(0) = 0$.
The general model ($q_j, g_i(x)$ with $g_i(0) > 0$) leads to the bound (3).
The transformed model ($q_j^{'i} = q_j(1-g_i(0)), g_i^{'}(x) = \frac{g_i(x)}{1-g_i(0)}$ with $g_i^{'}(0) = 0$) leads to the bound

$$\epsilon = \frac{1}{r+1} \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij} \frac{1-q_j(1-g_i(0))}{1-2q_j(1-g_i(0))} \left[\frac{q_j(1-g_i(0))}{1-q_j(1-g_i(0))}\right]^{r+1}$$

It is clear that the error bound of the transformed model will always be lower than the one of model 1. Thus, the conclusion of this section is : when performing the De Pril's approximation, it is always better to use a model where $g_i(0) = 0$.
Note that in that case the formulae are a little bit more complicated because the double sum

$$\sum_{i=1}^{a}\sum_{j=1}^{b}$$

in formula (2) has to be replaced by the sum

$$\sum_{i=1}^{a} \sum_{j=(i-1)b+1}^{ib}$$

4. The number of arithmetic operations with the De Pril's approximation

Dhaene and Vandebroek (1995) relied on the one stage formula in order to draw the conclusion given in section 2 about the number of multiplications needed to perform their exact and the De Pril's approximate algorithms.
However we can see that the De Pril's approximation is always better when it is considered as a double stage algorithm : $t^{(r)}(x)$ has to be evaluated from $x = 1$ to $x = r \max_i m_i$ which needs $2abr^2 \max_i m_i$ multiplications. This number does not depend on s and may be neglected.
The recursion needs $(2r \max_i m_i + 1)s$ multiplications.
Now the conclusion of this section is that the efficiency if far higher if one works with the double stage algorithm for the De Pril's approximation. It will always be better than the one stage formula :

$$(2r \max_i m_i + 1)s < (3ar + r(r+1) \sum_{i=1}^{a} (m_i - 1))s$$

It will be better than the exact Dhaene and Vandebroek (1995) formula if

$$r \max_i m_i < b \sum_{i=1}^{a} m_i$$

This will certainly be the case when $r < b$ and may be the case when $r > b$, precisely when

$$r < b \frac{\sum_{i=1}^{a} m_i}{\max_i m_i}$$

5. The bivariate individual risk model

From now on we will assume that the random variable giving the claim amount is bivariate. We will use the following definitions :

(X_{1i}, X_{2i})	:	random vector of claim amount
$g_i(x_1, x_2)$	:	probability function of (X_{1i}, X_{2i})
(m_{1i}, m_{2i})	:	maximum claim amounts for risk of type i
$G_i(u, v)$	:	probability generating function of the claim amounts
(S_1, S_2)	:	random vector of the aggregate claim amounts
$p(s_1, s_2)$	:	probability function of (S_1, S_2)
$P(u, v)$	:	probability generating function of (S_1, S_2)
q_j	:	probability of no claim for the risk of type j
$f(x) = O(x)$	:	$f(x)$ is of the same order as x

We will furthermore assume that (X_{1i}, X_{2i}) is defined on the non-negative integers and that $g_i(0, 0) = 0$ for the same reasons explained in section 3.
The reason why we extend the model to a bivariate setting will be described in the application in section 10.

In the following sections we will use the concept of bivariate ordinary generating functions. This is a generalization of the ordinary generating function (see Panjer and Willmot(1992) for an application in actuarial sciences). Multivariate ordinary generating functions are trivially extended.
Let us assume a sequence $\{a_{n,m}\ ,\ n \in \mathbb{Z}, m \in \mathbb{Z}\}$ of real numbers.
The ordinary generating function of this sequence is defined as

$$T_{a_{n,m}}(u, v) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_{n,m} u^n v^m$$

Of course u and v must be chosen in such a way that the sum exists.
Ordinary generating functions have the following nice properties :

- There is a one-to-one correspondence between $\{a_{n,m}\ ,\ n \in \mathbb{Z}, m \in \mathbb{Z}\}$ and $T_{a_{n,m}}(u, v)$
- $a_{n,m} = \frac{1}{n!m!} \frac{\partial^n \partial^m T_{a_{n,m}}(u,v)}{\partial u^n \partial v^n}|_{u=0,v=0}$
- $c_{n,m} = \alpha a_{n,m} + \beta b_{n,m} \Leftrightarrow T_{c_{n,m}}(u, v) = \alpha T_{a_{n,m}}(u, v) + \beta T_{b_{n,m}}(u, v)$
- $c_{n,m} = \sum_{k=0}^{n} \sum_{l=0}^{m} a_{k,l} b_{n-k,m-l} \Leftrightarrow T_{c_{n,m}}(u, v) = T_{a_{n,m}}(u, v) T_{b_{n,m}}(u, v)$
- $T_{na_{n,m}}(u, v) = u \frac{\partial}{\partial u} T_{a_n}(u, v)$

The philosophy behind using ordinary generating functions is the following :

- look for a relation between some sequences $a_{n,m}, b_{n,m}, c_{n,m}, \ldots$
- go in the (u, v) map where the calculations become easier (think of the convolution that becomes a product)

- go back in the initial map by inverting the expression in (u, v) thanks to the properties

In this paper, the sequences $a_{(n,m)}$ will be probability functions. As a consequence we will not have problems of convergence for the ordinary generating functions : $|u| < \infty$, $|v| < \infty$. In the present case, ordinary generating functions are just probability generating functions (pgf).
From now on we will only refer to pgf and we will use them extensively in the sense of ordinary generating functions .

6. Approximate formula

In this section we will extend the methodology developed in Dhaene and De Pril (1994).
We have

$$P(u,v) = \sum_{s_1=0}^{\infty}\sum_{s_2=0}^{\infty} p(s_1,s_2)u^{s_1}v^{s_2} \tag{4}$$

$$\ln P(u,v) = \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} t(x_1,x_2)u^{x_1}v^{x_2} \tag{5}$$

Note that as $p(0,0) > 0$, (5) is well defined in the neighbourhood of $(0,0)$.
As we immediately have

$$\begin{aligned}
\frac{\partial}{\partial u}P(u,v) &= P(u,v)\frac{\partial}{\partial u}\ln P(u,v) \\
\frac{\partial}{\partial v}P(u,v) &= P(u,v)\frac{\partial}{\partial v}\ln P(u,v)
\end{aligned}$$

we get, inverting the preceding expressions

$$p(0,0) = e^{t(0,0)}$$

$$s_1 p(s_1,s_2) = \sum_{x_1=1}^{s_1}\sum_{x_2=0}^{s_2} x_1 t(x_1,x_2)p(s_1-x_1,s_2-x_2) \quad s_1 \geq 1 \tag{6}$$

$$s_2 p(s_1,s_2) = \sum_{x_1=0}^{s_1}\sum_{x_2=1}^{s_2} x_2 t(x_1,x_2)p(s_1-x_1,s_2-x_2) \quad s_2 \geq 1 \tag{7}$$

In the particular case of the individual risk model, we have

$$P(u,v) = \prod_{i=1}^{a}\prod_{j=1}^{b}[p_j + q_j G_i(u,v)]^{n_{ij}}$$

An expansion of $\ln P(u,v)$ gives

$$\ln P(u,v) = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}\ln p_j + \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\left(\frac{q_i}{p_i}\right)^k G_i^k(u,v)$$

Inverting this expression gives

$$\begin{aligned} t(0,0) &= \sum_{i=1}^{n} \ln p_i \\ t(x_1,x_2) &= \sum_{i=1}^{a} \sum_{k=1}^{\max(x_1,x_2)} \sum_{j=1}^{b} n_{ij} \frac{(-1)^{k+1}}{k} \left(\frac{q_j}{p_j}\right)^k g_i^{*^k}(x_1,x_2) \end{aligned}$$

Calculating the $t(x_1,x_2)$ in a first stage and inserting these quantities in the bivariate recursion (6),(7) in a second stage will give us an exact algorithm.
As explained in section 4, we will always find the $t(x_1,x_2)$ in a first stage before inserting these values in the bivariate recursion. We certainly do not try to find a one stage bivariate recursion.
However this two-stage bivariate recursion is $O(s_1^2 s_2^2)$ and would be very time consuming. Nevertheless there is a trivial approximation for it which will be $O(s_1 s_2)$ consisting in truncating the sum in k to a certain integer r which will be very small most of the time.
Let $p^{(r)}(s_1,s_2)$ be the r^{th} order approximation for $p(s_1,s_2)$.
$t^{(r)}(x_1,x_2)$ be the r^{th} order approximation for $t(x_1,x_2)$.
We have

$$p^{(r)}(0,0) = e^{t^{(r)}(0,0)} \tag{8}$$

$$s_1 p^{(r)}(s_1,s_2) = \sum_{x_1=1}^{s_1} \sum_{x_2=0}^{s_2} x_1 t^{(r)}(x_1,x_2) p^{(r)}(s_1-x_1,s_2-x_2) \quad s_1 \geq 1 \tag{9}$$

$$s_2 p^{(r)}(s_1,s_2) = \sum_{x_1=0}^{s_1} \sum_{x_2=1}^{s_2} x_2 t^{(r)}(x_1,x_2) p^{(r)}(s_1-x_1,s_2-x_2) \quad s_2 \geq 1 \tag{10}$$

$$t^{(r)}(0,0) = \sum_{i=1}^{n} \ln p_i \tag{11}$$

$$t^{(r)}(x_1,x_2) = \sum_{i=1}^{a} \sum_{k=1}^{\min(r,\max(x_1,x_2))} \sum_{j=1}^{b} n_{ij} \frac{(-1)^{k+1}}{k} \left(\frac{q_j}{p_j}\right)^k g_i^{*^k}(x_1,x_2) \tag{12}$$

It is clear that our bivariate derivation trivially extends to a multivariate setting. As we will see in the next sections our algorithm written in the map of probability functions authorizes to have a look at the calculations time. This is not the case for the derivation of Dickson and Waters (1999) because (11)-(12) are related to the ordinary generating functions in their paper.
Note that the algorithm can also be discovered with the use of the multivariate De Pril transform as mentioned in Sundt (1999).

Using the same arguments as Dhaene and De Pril (1994), it is easy to extend their theorem 1 (see for example Sundt (1998):

Theorem 1 *If there exists a real number ϵ such that*

$$\sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} |t(x_1,x_2) - t^{(r)}(x_1,x_2)| \leq \epsilon \tag{13}$$

then the following error bound holds

$$\sum_{s_1=0}^{\infty}\sum_{s_2=0}^{\infty}|p(s_1,s_2)-p^{(r)}(s_1,s_2)|\le e^{\epsilon}-1$$

For our bivariate De Pril's approximation, it is also easy to show that the bound for (13) remains the same as in the univariate case :

$$\sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty}|t(x_1,x_2)-t^{(r)}(x_1,x_2)|\le \frac{1}{r+1}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\frac{p_j}{p_j-q_j}\left(\frac{q_j}{p_j}\right)^{r+1}$$

So the error bound for the $p(s_1,s_2)$ is controlled in the same way as in the univariate case.

7. Exact bivariate formula

In this section we extend the exact formula given by Dhaene and Vandebroek (1995) to the bivariate setting.

Theorem 2 *With the same notations as in the previous section, an exact solution for the individual risk model is given by*

$$p(0,0) = \prod_{i=1}^{a}\prod_{j=1}^{b}p_j^{n_{ij}} \tag{14}$$

$$s_1p(s_1,s_2) = \sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}v_{ij}(s_1,s_2) \quad s_1\ge 1\,,\ s_2\ge 0 \tag{15}$$

$$\begin{aligned} v_{ij}(s_1,s_2) &= \frac{q_j}{p_j}\sum_{x_1=0}^{m_{i1}}\sum_{x_2=0}^{m_{i2}}g_i(x_1,x_2)(x_1p(s_1-x_1,s_2-x_2)-v_{ij}(s_1-x_1,s_2-x_2)) \quad s_1\ge 1\,,\ s_2\ge 0 \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{16}$$

$$s_2p(s_1,s_2) = \sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}w_{ij}(s_1,s_2) \quad s_2\ge 1\,,\ s_1\ge 0 \tag{17}$$

$$\begin{aligned} w_{ij}(s_1,s_2) &= \frac{q_j}{p_j}\sum_{x_1=0}^{m_{i1}}\sum_{x_2=0}^{m_{i2}}g_i(x_1,x_2)(x_2p(s_1-x_1,s_2-x_2)-w_{ij}(s_1-x_1,s_2-x_2)) \quad s_2\ge 1\,,\ s_1\ge 0 \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{18}$$

Proof :

$$\begin{aligned} P(u,v) &= \sum_{s_1=0}^{\infty}\sum_{s_2=0}^{\infty}p(s_1,s_2)u^{s_1}v^{s_2} \\ &= \prod_{i=1}^{a}\prod_{j=1}^{b}[p_j+q_jG_i(u,v)]^{n_{ij}} \end{aligned}$$

$$\frac{\partial}{\partial u}P(u,v) = \sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}V_{ij}(u,v) \tag{19}$$

$$V_{ij}(u,v) = \sum_{s_1=0}^{\infty}\sum_{s_2=0}^{\infty}v_{ij}(s_1+1,s_2)u^{s_1}v^{s_2}$$

$$
\begin{aligned}
(20) \qquad\qquad & = \frac{q_j \frac{\partial}{\partial u} G_i(u,v) P(u,v)}{p_j + q_j G_i(u,v)} \\
\frac{\partial}{\partial v} P(u,v) &= \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} W_{ij}(u,v) \\
W_{ij}(u,v) &= \sum_{s_1=0}^{\infty} \sum_{s_2=0}^{\infty} w_{ij}(s_1, s_2+1) u^{s_1} v^{s_2} \\
&= \frac{q_j \frac{\partial}{\partial v} G_i(u,v) P(u,v)}{p_j + q_j G_i(u,v)}
\end{aligned}
$$

Multiplying both sides of (19) by u and inverting gives (15)
Multiplying both sides of (20) by u and rearranging gives

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

which gives (16) after inverting.
(17) and (18) are obtained similarly. □

Theorem 2 is valid only if $g_i(0,0) = 0$. If this is not the case, the best solution, as explained in section 3, is to change the model in order to have a new $g_i(0,0) = 0$.

8. Comparison of the exact and approximate algorithms

The cost for the exact algorithm is

$$2b(\sum_{i=1}^{a} m_{i1} m_{i2} + a) s_1 s_2$$

The cost for the approximate algorithm is

$$(2r^2 \max_i m_{i1} \max_i m_{i2} + 1) s_1 s_2$$

plus some initial work not depending on s_1, s_2 in order to construct the $t^{(r)}(x_1, x_2)$.
We can see immediately that the approximate formula will be more efficient than the exact one if

$$r^2 < b \frac{\sum_{i=1}^{a} m_{i1} m_{i2}}{\max_i m_{i1} \max_i m_{i2}}$$

Note that in this bivariate case the fraction may be > or < than 1.

9. The bivariate collective risk model as an approximation

As shown in De Pril and Dhaene (1992), the individual risk model can be approached by a collective risk model. They derive error bounds when working with the collective instead of the individual risk model. It is easy to see that their results remain in the bivariate setting. Namely, we have :

$$f^{CP}(x_1,x_2)=\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^k}{k!}g^{*k}(x_1,x_2)$$

with

$$\lambda=\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}q_j$$

$$g(x_1,x_2)=\frac{1}{\lambda}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}q_jg_i(x_1,x_2)$$

With this bivariate Compound Poisson approximation, we have the following results :

$$\begin{aligned}d_K((S_1^{ind},S_2^{ind}),(S_1^{CP},S_2^{CP}))&\le\frac{1}{2}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}q_j^2\\ d_{TV}((S_1^{ind},S_2^{ind}),(S_1^{CP},S_2^{CP}))&\le\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}q_j^2\end{aligned}$$

where d_K denotes the Kolmogorov distance while d_{TV} denotes the total variation distance :

$$\begin{aligned}d_K((X_1,X_2),(Y_1,Y_2))&=\sup_{(x_1,x_2)\in\mathbb{Z}^2}|\mathbb{P}(X_1\le x_1,X_2\le x_2)-\mathbb{P}(Y_1\le x_1,Y_2\le x_2)|\\ d_{TV}(X,Y)&=\sup_{A\subset\mathbb{Z}^2}|\mathbb{P}((X_1,X_2)\in A)-\mathbb{P}((Y_1,Y_2)\in A)|\end{aligned}$$

The bivariate Compound Poisson distribution is discussed in Walhin and Paris (2000) where a bivariate Panjer's algorithm is given in order to handle such distributions. Naturally, the multivariate extension is immediate.

10. Application : excess of loss reinsurance with reinstatements

In this section we will use the results described before in order to study the effect of an excess of loss cover with reinstatements on a fire portfolio. The methodology has been described in Walhin and Paris (2000) and we will now give an example based on the following data.

		f1	f2	f3	f4
	14				0,1
	12				0,1
	10			0,125	0,1
	8			0,125	0,1
	6			0,125	0,1
	5			0,125	0,1
	4		0,25	0,125	0,1
	3		0,25	0,125	0,1
	2	0,5	0,25	0,125	0,1
	1	0,5	0,25	0,125	0,1
qj	0,001	200	140	120	100
	0,002	170	140	120	100
	0,003	100	140	120	100

Table 2: Fire insurance portfolio

The loading of the cedent is 50%.
The loading of the reinsurer is 100%.
The premiums are calculated according to the expected value premium principle.
The reinsurer sells excess of loss cover $8\,xs\,6$.
The pure reinsurance premium is given by :

0 reinst. 1.61962	1 free reinst. 1.73527	2 free reinst. 1.73987	3 free reinst. 1.74
	1 paid reinst. at 100% 1.44311	2 paid reinst. at 100% 1.42975	3 paid reinst. at 100% 1.42917

Table 3: Reinsurance premiums

The fact that the premium for paid reinstatements reduces with the number of reinstatements is not illogical because it would be exceptional for the layer to be fully exhausted more than twice. So with two or more reinstatements, the ceding company has more capacity but it will pay more reinstatement premiums for a capacity that will most probably be unnecessary.

We are now interested in the expected gain and the adjustment coefficient of the cedent for different types of excess of loss covers. We find :

	0 reinst.	1 free reinst.	2 free reinst.	3 free reinst.
$\mathbb{E}G$	3.7653	3.6497	3.6451	3.645
R	0.1276	0.1468	0.1505	0.1508
		1 paid reinst. at 100%	2 paid reinst. at 100%	3 paid reinst. at 100%
$\mathbb{E}G$		3.64973	3.64513	3.645
R		0.1264	0.1232	0.1226

Table 4: Expected gain and adjustment coefficient in the individual risk model

We can make the following remarks :

- The covers with 2 or 3 paid reinstatements are clearly not interesting
- The free reinstatements are more interesting than the paid ones. This result has been established in general in Walhin and Paris (2000).
- Clearly a choice has to be made between 0 reinstatement which gives a higher expected gain but a poor adjustement coefficient and 1 free reinstatement that has the opposite properties.

Similar results are found using the Poisson approximation with the bivariate Panjer's algorithm described in Walhin and Paris (2000) :

	0 reinst.	1 free reinst.	2 free reinst.	3 free reinst.
$\mathbb{E}G$	3.7656	3.6499	3.6452	3.6451
R	0.1273	0.1469	0.1503	0.1504
		1 paid reinst. at 100%	2 paid reinst. at 100%	3 paid reinst. at 100%
$\mathbb{E}G$		3.6499	3.6452	3.6451
R		0.1263	0.1230	0.1224

Table 5: Expected gain and adjustment coefficient in the collective risk model

Note that in practice the calculation will be done with the reinsurance premiums given by the reinsurer and not with the ones calculated by ourselves with the expected value premium principle. As, most of the time, a lot of reinsurers make quotations, it can give arbitrage opportunities to the user of such models.

Of course multiple layers may be treated with the trivial multivariate extension of the individual risk model.

In order to make the calculations to find the adjustment coefficient, it was only necessary to compute the bivariate aggregate distribution of the portfolio for the risks f_3 and f_4 because the other risks never reach the layer. For them the moment generating function is sufficient in order to make the calculations.

		f3					f4								
		0	1	2	3	4	0	1	2	3	4	5	6	7	8
	14						0								
	12						0								
	10	0					0								
	8	0					0								
	6	0,125		0,125		0,125	0,1		0,1		0,1		0,1		0,1
	5	0,125					0,1								
	4	0,125					0,1								
	3	0,125					0,1								
	2	0,125					0,1								
	1	0,125					0,1								
qj	0,001	120					100								
	0,002	120					100								
	0,00	12					10								

Table 6: Part of the portfolio to be reinsured

In order to give an application of the result of section 9, we compare the bounds and the true errors for the Compound Poisson approximation of the individual risk model of table 6.
We find

$$\begin{aligned} d_{TV}((S_1^{ind}, S_2^{ind}), (S_1^{CP}, S_2^{CP})) = 0.0005942 &\quad \leq \quad \text{bound} = 0.00308 \\ d_K((S_1^{ind}, S_2^{ind}), (S_1^{CP}, S_2^{CP})) = 0.0004117 &\quad \leq \quad \text{bound} = 0.00154 \end{aligned}$$

We see that the approximation is very good. This is not a surprise as the q_j are very small.

The distribution of (S_1^{ind}, S_2^{ind}) has been evaluated with the exact bivariate algorithm of section 7 because it needs less multiplications than the De Pril's approximation unless $r = 2$.

11. Application : convolutions of bivariate distributions

In this section we consider a portfolio with n identical policies. ($p = p_j \quad \forall j$ and $f = f_i \quad \forall i$).
From theorem 2 we get

$$\begin{aligned} p(0,0) &= p^n \\ s_1 p(s_1,s_2) &= n v(s_1,s_2) \\ v(s_1,s_2) &= \frac{q}{p} \sum_{x_1=0}^{\min(s_1,m_1)} \sum_{x_2=0}^{\min(s_2,m_2)} f(x_1,x_2)(x_1 p(s_1-x_1,s_2-x_2) - v(s_1-x_1,s_2-x_2) \quad s_1 \geq 1 \\ s_2 p(s_1,s_2) &= n w(s_1,s_2) \\ w(s_1,s_2) &= \frac{q}{p} \sum_{x_1=0}^{\min(s_1,m_1)} \sum_{x_2=0}^{\min(s_2,m_2)} f(x_1,x_2)(x_2 p(s_1-x_1,s_2-x_2) - w(s_1-x_1,s_2-x_2) \quad s_2 \geq 1 \end{aligned}$$

which immediately gives

$$\begin{aligned}
p(0,0) &= p^n \\
s_1 p(s_1,s_2) &= \frac{q}{p} \sum_{x_1=0}^{\min(s_1,m_1)} \sum_{x_2=0}^{\min(s_2,m_2)} \left(\frac{(n+1)x_1 - s_1}{n} \right) f(x_1,x_2) p(s_1-x_1, s_2-x_2) \quad s_1 \geq 1 \\
s_2 p(s_1,s_2) &= \frac{q}{p} \sum_{x_1=0}^{\min(s_1,m_1)} \sum_{x_2=0}^{\min(s_2,m_2)} \left(\frac{(n+1)x_2 - s_2}{n} \right) f(x_1,x_2) p(s_1-x_1, s_2-x_2) \quad s_2 \geq 1
\end{aligned}$$

which is the bivariate extension of De Pril's (1985) formula for the convolutions of arithmetic distributions.
This formula has been deducted using other arguments in Walhin and Paris (2000) and in Sundt(1999).

12. Conclusion

We have noted in this paper why the De Pril's approximation has to be used as a two stage algorithm. Moreover, we have noted that it is always better to work with a model for which the claims distribution has no mass at the origin.
On the other hand we have extended the classical individual risk model to a bivariate setting. A natural application of this bivariate setting is the evaluation of the aggregate claims distribution of the ceding company when it buys an excess of loss cover with reinstatements. An exact algorithm is provided. It is the natural extension of the algorithm of Dhaene and Vandebroek (1995). An approximate algorithm is given. It is the natural extension of the De Pril's approximation.
A comparison between the bivariate individual model and its Compound Poisson approximation is also provided.
Naturally it is trivial to extend the models described in this paper to a multivariate setting.

References

References

D.C.M. Dickson and H.R. Waters (1999). Multi-Period Aggregate Loss Distributions for a Life Portofolio. *Astin Bulletin*, 29:295–309.

De Pril, N. (1985). Recursions for Convolutions of Arithmetic Distributions. *Astin Bulletin*, 15:135–139.

De Pril, N. (1986). On the Exact Computation of the Aggregate Claims distribution in the Individual Life Model. *Astin Bulletin*, 16:21–45.

De Pril, N. (1988). Improved Approximations for the Aggregate Claims Distributions of a Life Insurance Portfolio. *Scandinavian Actuarial Journal*, 1988:61–68.

De Pril, N. (1989). The Aggregate Claims Distribution in the Individual Model with Arbitrary Positive Claims. *Astin Bulletin*, 19:9–24.

De Pril, N. and Dhaene, J. (1992). Error Bounds for Compound Poisson Approximations of the Individual Risk Model. *Astin Bulletin*, 22:135–148.

Dhaene, J. and De Pril, N. (1994). On a Class of Approximative Computation Methods in the Individual Risk Model. *Insurance : Mathematics and Economics*, 14:181–196.

Dhaene, J. and Vandebroek, M. (1995). Recursions for the Individual Model. *Insurance : Mathematics and Economics*, 16:31–38.

Kornya, P. (1983). Distribution of Aggregate Claims in the Individual Model. *Transactions of the Society of Actuaries*, 35:823–836.

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

Panjer, H.H. and Willmot, G.E. (1992). *Insurance Risk Models.* Society of Actuaries.

Sundt, B. (1998). On Error Bounds for Multivariate Distributions. Research Paper 60, Centre for Actuarial Studies, University of Melbourne.

Sundt, B. (1999). Discussion on "D.C.M. Dickson and H.R. Waters" Multi-Period Aggregate Loss Distributions for a Life Portofolio. *Astin Bulletin*, 29:311–314.

Walhin, J.F. and Paris, J. (2000). Excess of Loss Reinsurance with Reinstatements and Ruin of the Cedent. Discussion Paper 0003, Institut de Statistique, UCL.

José Paris
Jean-François Walhin
Institut de Statistique
Voie du Roman Pays, 20
B-1348 Louvain-la-Neuve
Belgique
e-mail : walhin@stat.ucl.ac.be

Jean-François Walhin
Secura Belgian Re
Rue Montoyer, 12
B-1000 Bruxelles