MULTI-PERIOD AGGREGATE LOSS DISTRIBUTIONS
FOR A LIFE PORTFOLIO

By

David C.M. Dickson and Howard R. Waters

April 13, 1999

Abstract

Algorithms for the calculation of the distribution of the aggregate claims from a life insurance portfolio have been derived by Kornya (1983), Hipp (1986) and De Pril (1986 and 1989). All these authors considered the distribution of the aggregate claims over a single period. In this paper we derive algorithms for the calculation of the joint distribution of the aggregate claims from a life portfolio over several periods.

1. Introduction

Several authors, notably Kornya (1983), Hipp (1986) and De Pril (1986 and 1989) have derived algorithms for the calculation of the distribution of the aggregate claims from a life insurance portfolio. In each case, the author considers the aggregate claims over a single period of time. However, a distinguishing feature of most life insurance portfolios is that the term of the policies is greater than one year and, taking a year as a natural time period, the aggregate claims from year to year are likely to be correlated from the very nature of life insurance. With these features in mind, it is of interest to derive algorithms for the calculation of the joint distribution of aggregate claims over several time periods from a life insurance portfolio. In this paper we present such algorithms.

In the following section we define our notation, make some assumptions and set out our problem in precise terms. An algorithm for the solution to our problem is presented as Result 1 in Section 3. This result is a multi-period extension of De Pril's (1989) Theorem 1, and the proof follows along the same lines. In Section 4 we present a second algorithm, which can be regarded as an approximation to the (exact) algorithm in Section 3. An error bound for this approximate algorithm is given in Result 2. The approximation corresponds to De Pril's (1989) formulae (16.a) and (16.b) and the proof of our Result 2 is similar to that of this Theorem 3.

In Section 5 we describe a hypothetical portfolio of endowment assurances which we use to illustrate our algorithms. In particular, we are interested in the joint distribution of the losses in two successive years. We use the algorithms of Sections 3 and 4 to calculate this joint distribution exactly and approximately. We compare the results with those obtained from approximating this joint distribution in different ways.

2. Preliminaries

2.1. Set-up

We consider a life portfolio consisting of a total of N independent policies. For each policy j, $j = 1, ..., N$, we are interested in an m-dimensional vector of real-valued random variables, $\underline{Z}_j = (Z_{j1}, ..., Z_{jm})$, determined by the life history of the policy. For example:

1. Z_{jk} could represent the sum assured paid in year k in respect of policy j.
2. Z_{jk} could represent the present value of the loss in year k in respect of policy j in a portfolio of, for example, endowment assurances, or annuities or disability policies.
3. The time periods $k = 1, ..., m$ need not be disjoint; they could be overlapping time periods of unequal length.

The important feature is that for a given policy j, the random variables $\{Z_{jk}\}_{k=1}^{m}$ are *not* assumed to be independent. For convenience, we will refer to Z_{jk} as "the loss from policy j in year k".

We suppose that the policies can be grouped into n ($\leq N$) classes such that policies in the same class are probabilistically identical. More precisely, we assume that if policies j and l belong to the same class then $\underline{Z}_j$ and $\underline{Z}_l$ are identically distributed (as well as being independent when $j \neq l$). Let n_i denote the number of policies in class i, $i = 1, ..., n$, so that $N = \sum_{i=1}^{n} n_i$, and let $\underline{X}_i$ be a vector with the same distribution as $\underline{Z}_j$ for all policies in class i.

Throughout this paper, all vectors are of dimension m and a vector $\underline{z}$ has components $(z_1, z_2, ..., z_m)$. We will use the notation $\underline{y} \leq \underline{x}$ to mean that $y_k \leq x_k$ for $k = 1, ..., m$, and the notation $\underline{y} < \underline{x}$ to mean that $y_k \leq x_k$ for $k = 1, ..., m$, with at least one of the inequalities being strict.

2.2. Assumptions

1. For a policy from class i we assume that $\underline{X}_i$ is an integer-valued vector. In practical applications, this may require discretisation and hence approximation.

2. For a policy from class i and for $k = 1, ..., m$, let c_{ik} be the smallest possible value of X_{ik}. More precisely:

$$c_{ik} = \min\{x_{ik} | x_{ik} \in \mathbb{Z} \text{ and } \Pr(X_{ik} = x_{ik}) > 0\}$$

We denote $(c_{i1}, ..., c_{im})$ by $\underline{c}_i$. Let:

$$p_i = \Pr(\underline{X}_i = \underline{c}_i)$$
$$q_i = 1 - p_i$$

We assume that $0 < p_i < 1$.

2.3. Definitions

For $i = 1, 2, ..., n$, let:

$$\underline{Y}_i = \underline{X}_i - \underline{c}_i$$

and for $\underline{y} \neq \underline{0}$ let $h_i(\underline{y})$ be defined as:

$$h_i(\underline{y}) = \frac{1}{q_i} \Pr(\underline{Y}_i = \underline{y})$$

so that $h_i(\underline{y})$ is the probability that $\underline{X}_i = \underline{y} + \underline{c}_i$, given that $\underline{y} \neq \underline{0}$. (For convenience, we will define $h_i(\underline{0})$ to be 0.)

Let $B_i(\underline{s})$ be the probability generating function corresponding to the (multivariate) distribution determined by $h_i(\underline{y})$ so that:

$$B_i(\underline{s}) = \sum_{\underline{y} \neq \underline{0}} h_i(y_1, ..., y_m) s_1^{y_1} ... s_m^{y_m}$$

The following notation will be useful: for $j = 1, ..., m$, let $\underline{\delta}^{(j)}$ denote the vector whose j-th component is 1 and all of whose remaining components are 0.

2.4. Objective

Let $\underline{Y}$ be defined as follows:

$$\underline{Y} = \sum_{j=1}^{N} \underline{Z}_j - \sum_{i=1}^{n} n_i \underline{c}_i$$

so that the components of the vector $\underline{Y}$ are the total losses in successive years from the whole portfolio. Let $g(\underline{y})$ be the probability function for $\underline{Y}$, defined for all (non-negative) $\underline{y}$.

Our objective is to derive recursive formulae for the calculation of $g(\underline{y})$.

3. AN EXACT ALGORITHM

Result 1:

$$g(\underline{0}) = \prod_{i=1}^{n} (p_i)^{n_i} \tag{1}$$

$$g(y_1, \ldots, y_m) = \frac{1}{y_j} \sum_{\underline{0} \le \underline{x} \le \underline{y} - \underline{\delta}^{(j)}} w^{(j)}(\underline{x}) g(\underline{y} - \underline{\delta}^{(j)} - \underline{x}) \tag{2}$$

where $w^{(j)}(\underline{y}) = \sum_{i=1}^{n} w_i^{(j)}(\underline{y})$ and:

$$w_i^{(j)}(\underline{y}) = n_i \frac{q_i}{p_i} (y_j + 1) h_i(\underline{y} + \underline{\delta}^{(j)}) - \frac{q_i}{p_i} \sum_{\underline{0} \le \underline{x} \le \underline{y}} w_i^{(j)}(\underline{y} - \underline{x}) h_i(\underline{x}) \tag{3}$$

Proof: Formula (1) is immediate from the definitions of $\underline{Y}$ and p_i and from the independence of the policies.

Let $A(\underline{s})$ be the probability generating function of $\underline{Y}$. Then, from the independence of the policies, we have:

$$A(\underline{s}) = \prod_{i=1}^{n} (p_i + q_i B_i(\underline{s}))^{n_i} \tag{4}$$

Hence, for any j, $j = 1, \ldots, m$:

$$\frac{\partial}{\partial s_j} A(\underline{s}) = A(\underline{s}) \sum_{i=1}^{n} \frac{n_i q_i \frac{\partial}{\partial s_j} B_i(\underline{s})}{p_i + q_i B_i(\underline{s})}$$

Now define the functions $W_i^{(j)}(\underline{s})$ and $W^{(j)}(\underline{s})$ as follows:

$$W_i^{(j)}(\underline{s}) = \frac{n_i q_i \frac{\partial}{\partial s_j} B_i(\underline{s})}{p_i + q_i B_i(\underline{s})}$$

$$W^{(j)}(\underline{s}) = \sum_{i=1}^{n} W_i^{(j)}(\underline{s})$$

so that:

$$\frac{\partial}{\partial s_j} A(\underline{s}) = A(\underline{s}) W^{(j)}(\underline{s})$$

From their definitions, the definition of $B_i(\underline{s})$ and Assumption 1 in Section 2.2, it can be seen that both $W_i^{(j)}(\underline{s})$ and $W^{(j)}(\underline{s})$ can be written as power series in $s_1^{y_1} \ldots s_m^{y_m}$ for $\underline{y} \geq \underline{0}$. Hence we can write:

$$W_i^{(j)}(\underline{s}) = \sum_{\underline{y} \geq \underline{0}} w_i^{(j)}(\underline{y}) \, s_1^{y_1} \ldots s_m^{y_m}$$

$$W^{(j)}(\underline{s}) = \sum_{\underline{y} \geq \underline{0}} w^{(j)}(\underline{y}) \, s_1^{y_1} \ldots s_m^{y_m}$$

for some real-valued functions $w_i^{(j)}(\underline{y})$ and $w^{(j)}(\underline{y}) \left(= \sum_{i=1}^{n} w_i^{(j)}(\underline{y}) \right)$. Hence:

$$\sum_{\underline{x} \geq \underline{0}, x_j > 0} x_j g(\underline{x}) \, s_1^{x_1} \ldots s_j^{x_j - 1} \ldots s_m^{x_m}$$

$$= \left(\sum_{\underline{x} \geq \underline{0}} g(\underline{x}) \, s_1^{x_1} \ldots s_m^{x_m} \right) \left(\sum_{\underline{x} \geq \underline{0}} w^{(j)}(\underline{x}) \, s_1^{x_1} \ldots s_m^{x_m} \right) \qquad (5)$$

Now differentiate both sides of (5) with respect to s_1 (y_1 times), ..., s_j ($y_j - 1$ times), ..., s_m (y_m times) for any $\underline{y} \geq \underline{\delta}^{(j)}$ and put $s_1 = \ldots = s_m = 0$ to obtain:

$$\left(\prod_{k=1}^{m} y_k! \right) g(y_1, \ldots, y_m) = \frac{1}{y_j} \left(\prod_{k=1}^{m} y_k! \right) \sum_{\underline{0} \leq \underline{x} \leq \underline{y} - \underline{\delta}^{(j)}} w^{(j)}(\underline{x}) g(\underline{y} - \underline{\delta}^{(j)} - \underline{x})$$

This proves formula (2).
Now note that:

$$n_i q_i \frac{\partial}{\partial s_j} B_i(\underline{s}) = W_i^{(j)}(\underline{s})(p_i + q_i B_i(\underline{s}))$$

so that:

$$n_i q_i \sum_{\underline{x} \geq \underline{0}} x_j h_i(\underline{x}) \, s_1^{x_1} \ldots s_j^{x_j - 1} \ldots s_m^{x_m}$$

$$= \left(\sum_{\underline{x} \geq \underline{0}} w_i^{(j)}(\underline{x}) \, s_1^{x_1} \ldots s_m^{x_m} \right) \left(p_i + q_i \sum_{\underline{x} \neq \underline{0}} h_i(x_1, \ldots, x_m) \, s_1^{x_1} \ldots s_m^{x_m} \right) \qquad (6)$$

Let $\underline{y}$ be any non-negative vector. Formula (3) follows from differentiating (6) with respect to s_1 (y_1 times), ..., s_m (y_m times), putting $s_1 = \ldots = s_m = 0$, cancelling the term $\left(\prod_{k=1}^{m} y_k! \right)$ and rearranging.

Remark: Since formula (3) holds for any $\underline{y} \geq \underline{0}$, we have:

$$w_i^{(j)}(\underline{0}) = n_i \frac{q_i}{p_i} h_i(\underline{\delta}^{(j)}) \tag{7}$$

Starting from formula (7), formula (3) can be used recursively to calculate $w_i^{(j)}(\underline{y})$ for any $\underline{y} > 0$.

4. AN APPROXIMATE ALGORITHM

Throughout this section we will make the following extra assumption:

$$q_i < p_i \quad \text{for} \quad i = 1, ..., n \tag{8}$$

Recall that the random variable X_{ik} represents the "loss from policy i in year k". In many practical applications this loss will achieve its minimum value, i.e. c_{ik}, when no claim is made in year k and so p_i will be the probability that a policyholder from class i does not die in years $k = 1, 2, ..., m$. In such an application it is quite likely that assumption (8) will hold. This will be the case in the numerical example in the following section.

Using this assumption we will produce approximations to the joint probability function $g(\underline{y})$ and give bounds for the errors in these approximations.

First note that (4) can be written:

$$A(\underline{s}) = \exp(U(\underline{s}))$$

where:

$$U(\underline{s}) = \sum_{i=1}^{n} n_i \log(p_i + q_i B_i(\underline{s}))$$

$$= \sum_{i=1}^{n} n_i \left(\log(p_i) + \sum_{k=1}^{\infty} (-1)^{k+1} \left(\frac{q_i}{p_i} \right)^k \frac{1}{k} B_i(\underline{s})^k \right)$$

For each positive integer r, define $U(r; \underline{s})$ and $A(r; \underline{s})$ as follows:

$$U(r; \underline{s}) = \sum_{i=1}^{n} n_i \left(\log(p_i) + \sum_{k=1}^{r} (-1)^{k+1} \left(\frac{q_i}{p_i} \right)^k \frac{1}{k} B_i(\underline{s})^k \right)$$

$$A(r; \underline{s}) = \exp(U(r; \underline{s})) \tag{9}$$

We can write:

$$A(r; \underline{s}) = \sum_{\underline{y} \geq \underline{0}} g(r; \underline{y}) \, s_1^{y_1} \, ... \, s_m^{y_m}$$

for some function $g(r; \underline{y})$. We regard $g(r; \underline{y})$, or some simple adjustment of it, as an r-th order approximation to $g(\underline{y})$.

We can calculate $g(r; \underline{y})$ recursively using (9) as follows: for any j, $j = 1, ..., m$:

$$
\begin{aligned}
\frac{\partial}{\partial s_j} A(r; \underline{s}) &= \exp(U(r; \underline{s})) \frac{\partial}{\partial s_j} U(r; \underline{s}) \\
&= A(r; \underline{s}) \sum_{i=1}^{n} \sum_{k=1}^{r} n_i (-1)^{k+1} \left(\frac{q_i}{p_i}\right)^k B_i(\underline{s})^{k-1} \frac{\partial}{\partial s_j} B_i(\underline{s}) \\
&= A(r; \underline{s}) \sum_{\underline{y} \geq \underline{0}} u^{(j)}(r; \underline{y}) \, s_1^{y_1} ... s_m^{y_m}
\end{aligned}
\tag{10}
$$

for some function $u^{(j)}(r; \underline{y})$. Formula (10) corresponds to (5). We can derive the following formula in exactly the same way that (2) was derived from (5):

$$
g(r; x_1, ..., x_m) = \frac{1}{x_j} \sum_{\underline{0} \leq \underline{y} \leq \underline{x} - \underline{\delta}^{(j)}} u^{(j)}(r; \underline{y}) g(r; \underline{x} - \underline{\delta}^{(j)} - \underline{y})
\tag{11}
$$

The (possible) advantages of using $g(r; \underline{y})$ as the basis for an approximation to $g(\underline{y})$, i.e. using (11) rather than (2), are that values of $u^{(j)}(r; \underline{y})$ can be calculated directly from:

$$
\sum_{\underline{y} \geq \underline{0}} u^{(j)}(r; \underline{y}) \, s_1^{y_1} ... s_m^{y_m} = \sum_{i=1}^{n} \sum_{k=1}^{r} n_i (-1)^{k+1} \left(\frac{q_i}{p_i}\right)^k B_i(\underline{s})^{k-1} \frac{\partial}{\partial s_j} B_i(\underline{s})
\tag{12}
$$

whereas values of $w_i^{(j)}(\underline{y})$ have to be calculated recursively using (3) and (7), and also that if r is small and there are only a few vectors $\underline{y}$ for which $h_i(\underline{y})$ is non-zero, then there will only be a few vectors $\underline{y}$ for which $u^{(j)}(r; \underline{y})$ is non-zero. This should mean that fewer calculations are required to obtain values for $g(r; \underline{y})$ than are required to obtain values for $g(\underline{y})$.

Bound for the approximations are provided by the following result:

Result 2: For any $\underline{y} \geq \underline{0}$:

$$
\sum_{\underline{0} \leq \underline{x} \leq \underline{y}} |g(\underline{x}) - g(r; \underline{x})| \leq \exp\left(\sum_{i=1}^{n} \frac{n_i}{r+1} \left(\frac{q_i}{p_i}\right)^{r+1} \frac{p_i}{p_i - q_i}\right) - 1
$$

To prove this result we need the multi-dimensional analogue of the results on the partial ordering of power series used by Kornya (1983) and De pril (1988 and 1989). Let $C(\underline{s})$, $D(\underline{s})$ and $E(\underline{s})$ be any m-dimensional power series

defined for non-negative powers of $\underline{s}$, where $0 \leq s_j \leq 1$ for $j = 1, ..., m$, so that, for example:

$$C(\underline{s}) = \sum_{\underline{y} \geq \underline{0}} c(\underline{y}) \, s_1^{y_1} \, ... \, s_m^{y_m}$$

for some real valued function $c(\underline{y})$ defined for vectors $\underline{y} \geq \underline{0}$.

Definitions: We denote by $|C(\underline{s})|_{\underline{s}}$ the power series $\sum_{\underline{y} \geq \underline{0}} \left| c(\underline{y}) \right| s_1^{y_1} \, ... \, s_m^{y_m}$. We write:

$$C(\underline{s}) \leq_{\underline{s}} D(\underline{s})$$

whenever for any $\underline{y} \geq \underline{0}$ we have:

$$\sum_{\underline{0} \leq \underline{x} \leq \underline{y}} c(\underline{x}) \leq \sum_{\underline{0} \leq \underline{x} \leq \underline{y}} d(\underline{x})$$

The (one-dimensional version of the) following lemma appears in De Pril (1988 and 1989) and is proved by Kornya (1983). The proof is similar in the multi-dimensional case.

Lemma:
(i) $|C(\underline{s})|_{\underline{s}} \leq_{\underline{s}} \sum_{\underline{y} \geq \underline{0}} \left| c(\underline{y}) \right|$ whenever $\sum_{\underline{y} \geq \underline{0}} \left| c(\underline{y}) \right| < \infty$
(ii) $|C(\underline{s}) + D(\underline{s})|_{\underline{s}} \leq_{\underline{s}} |C(\underline{s})|_{\underline{s}} + |D(\underline{s})|_{\underline{s}}$
(iii) $|C(\underline{s}) \cdot D(\underline{s})|_{\underline{s}} \leq_{\underline{s}} |C(\underline{s})|_{\underline{s}} \cdot |D(\underline{s})|_{\underline{s}}$
(iv) $|\exp(C(\underline{s})) - 1|_{\underline{s}} \leq_{\underline{s}} \exp\left(|C(\underline{s})|_{\underline{s}}\right) - 1$
(v) $|C(\underline{s})|_{\underline{s}} \leq_{\underline{s}} |D(\underline{s})|_{\underline{s}} \Rightarrow |C(\underline{s})|_{\underline{s}} \cdot |E(\underline{s})|_{\underline{s}} \leq_{\underline{s}} |D(\underline{s})|_{\underline{s}} \cdot |E(\underline{s})|_{\underline{s}}$
(vi) $|C(\underline{s})|_{\underline{s}} \leq_{\underline{s}} |D(\underline{s})|_{\underline{s}} \Rightarrow \exp\left(|C(\underline{s})|_{\underline{s}}\right) - 1 \leq_{\underline{s}} \exp\left(|D(\underline{s})|_{\underline{s}}\right) - 1$

Proof of Result 2: Since:

$$A(\underline{s}) - A(r; \underline{s}) = A(\underline{s}) \cdot (1 - \exp(\log A(r; \underline{s}) - \log A(\underline{s})))$$

we can use part (iii) of the Lemma to write:

$$|A(\underline{s}) - A(r; \underline{s})|_{\underline{s}} \leq_{\underline{s}} |A(\underline{s})|_{\underline{s}} \cdot |\exp(\log A(r; \underline{s}) - \log A(\underline{s})) - 1|_{\underline{s}}$$

But:

$$|A(\underline{s})|_{\underline{s}} = \sum_{\underline{y} \geq \underline{0}} \left| g(\underline{y}) \right| s_1^{y_1} \, ... \, s_m^{y_m}$$

$$\leq_{\underline{s}} \sum_{\underline{y} \geq \underline{0}} \left| g(\underline{y}) \right| \leq_{\underline{s}} 1$$

using part (i) of the Lemma. Hence:

$$|A(\underline{s}) - A(r;\underline{s})|_{\underline{s}} \leq_{\underline{s}} |\exp(\log A(r;\underline{s}) - \log A(\underline{s})) - 1|_{\underline{s}}$$

$$\leq_{\underline{s}} \exp\left(|\log A(r;\underline{s}) - \log A(\underline{s})|_{\underline{s}}\right) - 1$$

$$= \exp\left(|U(r;\underline{s}) - U(\underline{s})|_{\underline{s}}\right) - 1$$

Also

$$|U(r;\underline{s}) - U(\underline{s})|_{\underline{s}} = \left| \sum_{i=1}^{n} \sum_{k=r+1}^{\infty} n_i(-1)^{k+1} \left(\frac{q_i}{p_i}\right)^{k} \frac{1}{k} B_i(\underline{s})^{k} \right|_{\underline{s}}$$

$$\leq_{\underline{s}} \sum_{i=1}^{n} \sum_{k=r+1}^{\infty} n_i \left(\frac{q_i}{p_i}\right)^{k} \frac{1}{k}$$

Hence, using part (vi) of the Lemma:

$$|A(\underline{s}) - A(r;\underline{s})|_{\underline{s}} \leq_{\underline{s}} \exp\left(\sum_{i=1}^{n} \sum_{k=r+1}^{\infty} n_i \left(\frac{q_i}{p_i}\right)^{k} \frac{1}{k}\right) - 1$$

$$\leq_{\underline{s}} \exp\left(\sum_{i=1}^{n} \frac{n_i}{r+1} \left(\frac{q_i}{p_i}\right)^{r+1} \frac{p_i}{p_i - q_i}\right) - 1$$

Hence, for any $\underline{y} \geq \underline{0}$

$$\sum_{\underline{0} \leq \underline{x} \leq \underline{y}} |g(\underline{x}) - g(r;\underline{x})| \leq \exp\left(\sum_{i=1}^{n} \frac{n_i}{r+1} \left(\frac{q_i}{p_i}\right)^{r+1} \frac{p_i}{p_i - q_i}\right) - 1$$

Remarks: Being a probability, $g(\underline{x})$ is always non-negative. However, there is no guarantee that $g(r;\underline{x})$ will always be non-negative. To avoid the embarrassment of approximating a probability by a negative number, we could use either $|g(r;\underline{x})|$ or $\max(g(r;\underline{x}), 0)$ as an r-th order approximation to $g(\underline{x})$. Note that the upper bound in Result 2 applies to these approximations since

$$\sum_{\underline{0} \leq \underline{x} \leq \underline{y}} |g(\underline{x}) - |g(r;\underline{x})|| \leq \sum_{\underline{0} \leq \underline{x} \leq \underline{y}} |g(\underline{x}) - g(r;\underline{x})|$$

and

$$\sum_{0 \le \underline{x} \le \underline{y}} |g(\underline{x}) - \max(g(r; \underline{x}), 0)| \le \sum_{0 \le \underline{x} \le \underline{y}} |g(\underline{x}) - g(r; \underline{x})|$$

Set $j = 1$ in (12), so that we are in the one-dimensional case considered by De Pril (1986 and 1989). The value of $u^{(j)}(r; x)$, for any $x \ge 0$, can be calculated from (12) by differentiating the right hand side x times with respect to s and then setting $s = 0$. Let $h_i^{*k}(.)$ denote the k-fold convolution of $h_i(.)$ with itself. Then $B_i(s)^k$ is the probability generating function of $h_i^{*k}(.)$ and it can be seen that:

$$u^{(j)}(r; x) = \sum_{i=1}^{n} \sum_{k=1}^{r} n_i(-1)^{k+1} \left(\frac{q_i}{p_i}\right)^k \left(\frac{1}{x}\right) h^{*k}(x+1)$$

A little algebra then shows that the approximation in this section corresponds to De Pril's approximation (1989, Section 4).

5. NUMERICAL EXAMPLE

5.1. The portfolio

Consider a portfolio of identical 10-year non-profit endowment assurances. The sum assured is payable at the end of the year of death, or on maturity, and premiums are payable annually in advance. The policyholders are all males and are aged 45 when their policies are effected. We are interested in the experience of this portfolio, including new entrants next year, over the next two years. If a policyholder survives for a year, or surrenders his policy at the end of a year, the loss to the insurer is zero. If a policyholder dies in a year, the loss to the insurer depends on the policyholder's age at the start of the year: values for these losses are shown in Table 1. These figures are broadly consistent with the losses from a policy with sum assured 16 and an annual premium of 1, where loss is defined as the sum assured payable minus the reserve being held. Also shown in Table 1 are the probability of death in the coming year for each age (these are taken from the AM(80) Select table, see CMI Report Number 10 (1990)) and the current number of policyholders at each age as a multiple of some positive integer K. One year from now, $3K$ policyholders, then aged 45, will enter the portfolio. Finally, for each policyholder who survives to age 46 there is a probability of 0.3 that he will surrender his policy before paying the premium then due.

TABLE 1

PORTFOLIO DETAILS

Age at start of the year	Loss on death in the year	Prob. of death in the year $\times 10^3$	Current number of policyholders
45	15	1.467	$3K$
46	14	2.064	$2K$
47	12	2.660	$2K$
48	11	3.003	$2K$
49	10	3.386	$2K$
50	8	3.813	$2K$
51	6	4.290	$2K$
52	4	4.821	$2K$
53	2	5.410	$2K$
54	0	6.065	$2K$

We want to calculate the joint distribution of (X_1, X_2), where X_1 and X_2 are the aggregate losses from this portfolio this year and next year, respectively. In the terminology of Section 2, we have:

$N = 16K$: note that those policyholders currently aged 54 will not contribute to either X_1 or X_2 and so can be ignored.

$n = 10$: the ten classes of policyholders are one for the new entrants next year $(i = 1)$ and one for each of the current ages 45 to 53 $(i = 2, ..., 10)$.

$n_i = 3K$ for $i = 1, 2$ and $n_i = 2K$ for $i = 3, ..., 10$.

$m = 2$ since we are considering the joint distribution of the losses over two time periods.

X_{ik} is the loss in year k from a policy in class i, so that, for example:

$$X_{3,1} = \begin{cases} 0 & \text{w.p. } 0.997936 \\ 14 & \text{w.p. } 0.002064 \end{cases}$$

$$X_{3,2} = \begin{cases} 0 & \text{w.p. } 0.002064 + 0.997936 \times 0.997340 = 0.997345 \\ 12 & \text{w.p. } 0.997936 \times 0.002660 = 0.002655 \end{cases}$$

$c_{ik} = 0$ for all i and k since for each policyholder there is either a positive (or zero) loss on death or a zero loss on survival or withdrawal.

Straightforward calculations give the values for the moments of X_1 and X_2, as multiples K, shown in Table 2, where $\rho(X_1, X_2)$ is the coefficient of correlation between X_1 and X_2.

TABLE 2

MOMENTS OF X_1 AND X_2

$E[X_1]$	$= 0.494129 \times K$
$V[X_1]$	$= 4.950015 \times K$
$E[X_2]$	$= 0.521714 \times K$
$V[X_2]$	$= 5.342768 \times K$
$\rho(X_1, X_2)$	$= -0.002549$

5.2. Numerical results

We calculated exact and approximate values of $\Pr(X_1 \leq x_1, X_2 \leq x_2)$ for $0 \leq x_1 \leq 200, 0 \leq x_2 \leq 200$ for both $K = 5$ and $K = 10$. For ease of presentation, we show in Tables 3a and 3b the results for $(x_1, x_2) = (x, x)$, where $x = 0, 5, ..., 25$, only. The key to these two tables is as follows:

(1) denotes the exact value of the probability calculated using Result 1 in Section 3. We also calculated these probabilities using the approximate method derived in Section 4 with $r = 3$; since the upper bound for the error is $\exp(K \times 1.5 \times 10^{-8}) - 1$, the results were the same as the exact values to at least four decimal places in both tables. Run times for the approximate algorithm are faster than those for the exact algorithm, but the difference is apparent only for large values of the argument (x_1, x_2). None of the calculated values of $g(r; \underline{x})$ was negative in our examples.

(2) denotes an approximation based on the assumption that X_1 and X_2 have independent normal distributions with means and variances as in Table 2.

(3) denotes an approximation based on the assumption that (X_1, X_2) has a bivariate normal distribution with moments as shown in Table 2.

(4) denotes an approximation based on the assumption that X_1 and X_2 are independent. The distributions of X_1 and X_2 have been calculated exactly using De Pril's (1989) one-dimensional recursive algorithm.

(5) denotes an approximation based on a method proposed by Wang (1997, formula (11.2)). This method produces an approximation to a joint distribution which has the correct marginal distributions for X_1 and X_2 and the correct covariance between the two distributions. A feature of this approximation is that the approximating function need not be a probability distribution; in particular, it need not always be non-negative and it need not be bounded above by one.

TABLE 3a

EXACT AND APPROXIMATE VALUES OF $\Pr(X_1 \le x, X_2 \le x) : K = 5$

x	(1)	(2)	(3)	(4)	(5)
0	0.5298	0.0951	0.0947	0.5304	0.5300
5	0.6460	0.4710	0.4706	0.6464	0.6461
10	0.8015	0.8635	0.8635	0.8017	0.8015
15	0.9554	0.9859	0.9859	0.9554	0.9554
20	0.9789	0.9994	0.9994	0.9789	0.9789
25	0.9926	1.0000	1.0000	0.9926	0.9926
30	0.9984	1.0000	1.0000	0.9984	0.9984

TABLE 3b

EXACT AND APPROXIMATE VALUES OF $\Pr(X_1 \le x, X_2 \le x) : K = 10$

x	(1)	(2)	(3)	(4)	(5)
0	0.2807	0.0573	0.0571	0.2813	0.2810
5	0.4119	0.2457	0.2453	0.4124	0.4121
10	0.6189	0.5680	0.5678	0.6192	0.6190
15	0.8486	0.8401	0.8401	0.8487	0.8485
20	0.9196	0.9626	0.9626	0.9197	0.9196
25	0.9650	0.9944	0.9944	0.9650	0.9650
30	0.9878	0.9995	0.9995	0.9878	0.9878

The numerical example in this section has been included to illustrate the range of alternative methods available for approximating the joint distribution of losses from a life portfolio and to test their accuracy when applied to a particular example. We cannot draw general conclusions from our numerical results but we can make the following points about Tables 3a and 3b:

(a) The numerical results for the independent normal and the bivariate normal approximations, columns (2) and (3), are very close to each other. This is presumably because the correlation between X_1 and X_2 is very small.

(b) This small correlation also explains why Wang's approximation, column (5), does not produce significantly different results from the approximation assuming independence of X_1 and X_2, column (4).

(c) The two normal approximations, columns (2) and (3), produce significantly poorer approximations than the approximation assuming independence and Wang's approximation, columns (4) and (5). As

K, the multiplier for the number of policies, increases, we would expect, by appealing to the Central Limit Theorem, that the two normal approximations would perform better. However, note that for K as low as 10, we have a portfolio with over 200 independent policies.

6. Acknowledgements

Note added in proof stage: Our aim in this paper has been to extend to the multivariate case the results (in the univariate case) derived by De Pril (1986 and 1989). Our methods follow closely those of De Pril. After submitting this paper to ASTIN we were made aware of the relationship between our results and those of Sundt (1998a and 1998b). These relationships are detailed in the Discussion paper by Sundt in this issue. We readily acknowledge that Sundt's work predates the work for our paper.

References

Continuous Mortality Investigation Reports, Number 10 (1990) Published by the Faculty of Actuaries and the Institute of Actuaries, Edinburgh and London.

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* 1988, 61-68.

DE PRIL, N. (1989) The aggregate claims distribution in the individual model with arbitrary positive claims. *ASTIN Bulletin* 19, 9-24.

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

KORNYA, P.S. (1983) Distribution of aggregate claims in the individual risk theory model. *Transactions of the Society of Actuaries* 35, 823-836.

SUNDT, B. (1998a) The multivariate De Pril transform. Center for Actuarial Studies, University of Melbourne, Research Paper Number 50.

SUNDT, B. (1998b) On error bounds for multivariate distributions. Centre for Actuarial Studies, University of Melbourne, Research Paper Number 60.

WANG, S.S. (1997) Aggregation of correlated risk portfolios: Models and algorithms. Institute of Insurance and Pension Research, University of Waterloo, Research Report 97-14.

DAVID C.M. DICKSON
Centre for Actuarial Studies
Faculty of Economics and Commerce
University of Melbourne
Parkville
Victoria 3052
Australia
email:D.Dickson@ecomfac.unimelb.edu.au

HOWARD R. WATERS
Department of Actuarial Mathematics and Statistics
Heriot-Watt University
Edinburgh
EH14 4AS
Scotland
email:H.R.Waters@ma.hw.ac.uk