On a Class of Approximative
Computation Methods
in the Individual Risk Model

by

Jan DHAENE

Nelson DE PRIL

CONTENTS

On a Class of Approximative Computation Methods in the Individual Risk Model

Jan DHAENE and Nelson DE PRIL

CATHOLIC UNIVERSITY OF LEUVEN
Belgium

Abstract

Exact calculations in the individual risk model are possible, but are very time consuming. Therefore, a number of recursive methods for approximate computation of the aggregate claims distribution and stop-loss premiums have been developed.
In the present paper a general class of such approximation methods is considered, containing the higher order approximations suggested by Kornya, Hipp and De Pril. In this way, the treatment of the different methods is unified and extended to a more general setting. Some new theoretical error bounds are derived, giving a quantitative measure of the accuracy of the approximations.

Keywords

1.   Introduction

Consider a random variable S representing the aggregate claims of an insurance portfolio in a certain reference period. Suppose that S has a finite mean, that the probability of no claims is strictly positive and that the individual claim amounts are positive and integer multiples of some convenient monetary unit, so that one can consider S as being defined on the non-negative integers.

Assume that an explicit expression for the probability generating function of S

$$P(u) = \sum_{s=0}^{\infty} p(s)u^s \qquad (1)$$

is available, but that the corresponding probabilities $p(s)$ are not known. The problem under consideration is to calculate the probabilities $p(s)$ numerically from the given expression $P(u)$.

In this paper the computation of the probabilities $p(s)$ will be based on a recursive scheme. To set up this scheme, coefficients $t(x)$ are introduced, which are defined by

$$\ln P(u) = \sum_{x=0}^{\infty} t(x)u^x \qquad (2)$$

Note that the assumption $p(0) > 0$ implies that $\ln P(u)$ is defined in a neighbourhood of zero, so that (2) is meaningful. Equalities between functions of power series will often be considered in the sequel of the paper. They always have to be interpreted as holding in some non-trivial interval of convergence.

Taking the derivative of both sides of (2) and equating the coefficients of the same power of u yields the following recursion for the probabilities $p(s)$

$$
\begin{cases}
p(0) = e^{t(0)} & \qquad\qquad\qquad\qquad\qquad (3.a) \\
\\
s\, p(s) = \sum_{x=1}^{s} x\, t(x)p(s-x) & \qquad s = 1, 2, \dots \qquad\qquad (3.b)
\end{cases}
$$

In this way the problem of computing the probabilities $p(s)$ is replaced by the problem of determining the appropriate coefficients $t(x)$. At first sight, one might think that the problem has only been shifted. However, in some important applications it is possible to give an explicit expression for the $t(x)$.

EXAMPLE 1. The compound Poisson model

In the classical collective risk model S is a compound Poisson variable, i.e.

$$P(u) = \exp\left\{\lambda \left[G(u) - 1\right]\right\} \tag{4}$$

with λ the Poisson parameter of the counting variable and $G(u)$ the probability generating function of the individual claim amounts

$$G(u) = \sum_{x=1}^{\infty} g(x) u^x \tag{5}$$

In this case, one immediately has that

$$\begin{cases} t(0) = -\lambda & (6.a) \\ t(x) = \lambda\, g(x) & x = 1, 2, \ldots \quad (6.b) \end{cases}$$

Inserting these values for $t(x)$ in *(3)* leads to the well-known recursive formula for compound Poisson distributions, see PANJER (1981).

$$\begin{cases} p(0) = e^{-\lambda} & (7.a) \\ s\, p(s) = \lambda \sum_{x=1}^{s} x\, g(x)\, p(s-x) & s = 1, 2, \ldots \quad (7.b) \end{cases}$$

EXAMPLE 2. The individual risk model

Consider a portfolio of n independent policies, labelled from 1 to n. Let $0 < p_i < 1$ be the probability that policy i produces no claims in the reference period and $q_i = 1 - p_i$ be the probability that this policy leads to at least one positive claim. Further, denote by

$$G_i(u) = \sum_{x=1}^{\infty} g_i(x)\, u^x \qquad i = 1, 2, \ldots, n \tag{8}$$

the probability generating function of the total claim amount of policy i, given that this policy has at least one claim. The probability generating function of the aggregate claims S of the portfolio is then given by

$$P(u) = \prod_{i=1}^{n} \left[p_i + q_i\, G_i(u) \right] \tag{9}$$

Now, $\ln P(u)$ can be expanded as

$$\ln P(u) = \sum_{i=1}^{n}\left[\ln p_i + \ln\left[1+\frac{q_i}{p_i}G_i(u)\right]\right]$$

$$= \sum_{i=1}^{n}\ln p_i + \sum_{i=1}^{n}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\left(\frac{q_i}{p_i}\right)^k G_i^k(u) \quad,$$

so that it can be written in the form (2), with coefficients $t(x)$ given by

$$\begin{cases} t(0) = \sum_{i=1}^{n}\ln p_i & (10.a) \\ t(x) = \sum_{k=1}^{x}\frac{(-1)^{k+1}}{k}\sum_{i=1}^{n}\left(\frac{q_i}{p_i}\right)^k g_i^{*k}(x) & x=1,2,\ldots\,, \quad (10.b) \end{cases}$$

where $g_i^{*k}(x)$ denotes the k-fold convolution of $g_i(x)$. Insertion of the $t(x)$-values in (3) leads to an exact recursive formula for the evaluation of the aggregate claims distribution in the individual model.

$$\begin{cases} p(0) = \prod_{i=1}^{n}p_i & (11.a) \\ s\,p(s) = \sum_{x=1}^{s}x\sum_{k=1}^{x}\frac{(-1)^{k+1}}{k}\sum_{i=1}^{n}\left(\frac{q_i}{p_i}\right)^k g_i^{*k}(x)\,p(s-x) & s=1,2,\ldots \quad (11.b) \end{cases}$$

This recursion was developed in two stages. The first stage is given in DE PRIL (1986), where the special case of the individual life model is considered. In this model each policy can only have one claim with a fixed amount. Then, the recursion (11) is simplified, since $g_i^{*k}(x) = \delta_{kc_i,x}$ with c_i the amount at risk of policy i and $\delta_{a,b}$ the Kronecker symbol, i.e. $\delta_{a,a}=1$ and $\delta_{a,b}=0$ if $a\neq b$. Numerical computations by KUON, REICH and REIMERS (1987) indicate that although the recursion formula is a much more powerful tool than the technique of straightforward convolutions, still a lot of computer time is required to calculate the distribution of S for a real life insurance portfolio.
The generalization to arbitrary positive claims is made in DE PRIL (1989). In this case the computation of the coefficients $t(x)$ is even more time consuming than for the individual life model, especially if the number of policies is large and if the conditional claim amount distributions $g_i(x)$ are defined for more than a few values.

The examples above show that the problem of computing the probabilities $p(s)$ numerically is not always solved satisfactory by deriving an explicit expression for the coefficients $t(x)$. This will only be the case if the $t(x)$ can easily be obtained, as in Example 1.

In case the $t(x)$ are too difficult to work with, it is of interest to consider approximative methods. Following the ideas of DE PRIL (1988 and 1989), approximations $f(s)$ of the probabilities $p(s)$ are set up by replacing the exact coefficients $t(x)$ in the recursion *(3)* by easy calculable values $h(x)$.

For a given choice of the sequence $\{h(x): x = 0, 1, ... \}$ approximations $f(s)$ of $p(s)$ for $s = 0, 1, ...$, are thus determined by the recursion

$$
\begin{cases}
f(0) = e^{h(0)} & \text{(12.a)} \\
s\, f(s) = \sum_{x=1}^{s} x\, h(x)\, f(s-x) & s = 1, 2, ... \quad \text{(12.b)}
\end{cases}
$$

Remark that $f(0)$ will always be positive, but for $s = 1, 2, ...$ the approximations $f(s)$ can be any value, positive or negative, depending on the choice of the coefficients $h(x)$.

Clearly a minimal requirement is that this choice leads to a sequence $\{f(s): s = 0, 1, 2, ...\}$ which has a generating function

$$
F(u) = \sum_{s=0}^{\infty} f(s)\, u^{s} \tag{13}
$$

with a strictly positive radius of convergence. In order for this to be true, it is assumed that $\sum_{x=0}^{\infty} h(x)\, u^{x}$ has a stricltly positive radius of convergence. In view of *(12)*, this implies that $F(u)$ can be written as

$$
F(u) = \exp\left(\sum_{x=0}^{\infty} h(x)\, u^{x} \right) \tag{14}
$$

For further reference, note that one also has

$$
F'(u) = F(u) \sum_{x=1}^{\infty} x\, h(x)\, u^{x-1} \tag{15}
$$

In actual applications more restrictive conditions will be imposed on the choice of the sequence $\{h(x): x = 0, 1, ...\}$, which will guarantee e.g. that $F(1) = \sum_{s=0}^{\infty} f(s)$ and $F'(1) = \sum_{s=1}^{\infty} s\, f(s)$ are finite.

In the following sections the quality of different approximations $f(s)$ will be discussed. This will be done by examining how these approximations perform with respect to the following two criteria, which should be satisfied simultaneously.

    i)   The time needed for computing the approximated aggregate claims distribution and related quantities, such as stop-loss premiums, should be as short as possible.

    ii)   The approximated values for the distribution function and stop-loss premiums should be as close as possible to the exact values.

The important point is that these two conditions should be satisfied at the same time. Indeed, it is not difficult at all to construct approximations which perform very good for one of the two criteria, but very bad for the other. The objective is thus to find, for the individual model, simple approximations $h(x)$ of $t(x)$ which lead to good approximations $f(s)$ of $p(s)$.

In Section 2 theoretical error bounds are derived which give a quantitative measure to assess the quality of $f(s)$ as approximation of $p(s)$. Section 3 shows how approximated stop-loss premiums can be calculated and gives bounds for the difference between the exact and approximated values. Applications of these results are given in Section 4, where the approximations proposed by KORNYA (1983), HIPP (1986) and DE PRIL (1988 and 1989) are examined and compared.

2.   Error Bounds for the Aggregate Claims Distribution

The following theorem gives an error bound for the difference between the probabilities $p(s)$ and their approximated values $f(s)$, corresponding with a given choice of the sequence $\{h(x):x=0,1,...\}$.

THEOREM 1. If there exists a real number ε such that

$$\sum_{x=0}^{\infty} |t(x)-h(x)| \le \varepsilon \qquad (16)$$

then the following error bound holds

$$\sum_{s=0}^{\infty} |p(s)-f(s)| \le e^{\varepsilon} - 1 \quad . \qquad (17)$$

Proof. In view of *(2)* and *(14)*, $F(u)$ can be written as

$$F(u) = A(u) P(u) \qquad (18)$$

with

$$A(u) = \sum_{x=0}^{\infty} a(x) u^x = \exp\left[\sum_{x=0}^{\infty} [h(x)-t(x)] u^x\right] \quad . \qquad (19)$$

Equating the coefficients of the same power of u in *(18)* gives

$$f(s) = \sum_{x=0}^{s} a(x) p(s-x) \qquad\qquad s=0,1,... \qquad (20)$$

from which it follows that

$$\sum_{s=0}^{\infty} |p(s) - f(s)| \leq |1 - a(0)| + \sum_{s=1}^{\infty} \sum_{x=1}^{s} |a(x)| \, p(s - x)$$

$$\leq |1 - a(0)| + \sum_{x=1}^{\infty} |a(x)| \quad . \tag{21}$$

From *(19)* the following recursion for the $a(x)$ can be obtained

$$\begin{cases} a(0) = e^{h(0) - t(0)} & \tag{22.a} \\[2mm] x \, a(x) = \sum_{y=1}^{x} y [h(y) - t(y)] \, a(x - y) & x = 1,2,\ldots \quad . \tag{22.b} \end{cases}$$

Now it is easy to prove by induction that

$$|a(x)| \leq b(x) \exp\left[\sum_{y=0}^{\infty} |t(y) - h(y)| \right] \qquad x = 0,1,\ldots \quad , \tag{23}$$

with $\{b(x) : x = 0,1,\ldots\}$ a compound Poisson probability function defined by

$$\begin{cases} b(0) = \exp\left[-\sum_{y=1}^{\infty} |t(y) - h(y)| \right] \\[3mm] x \, b(x) = \sum_{y=1}^{x} y \, |t(y) - h(y)| \, b(x - y) & x = 1,2,\ldots \quad . \end{cases}$$

Using *(16)* and *(23)* one gets

$$\sum_{x=1}^{\infty} |a(x)| \leq [1 - b(0)] \exp\left[\sum_{y=0}^{\infty} |t(y) - h(y)| \right]$$

$$\leq e^{\varepsilon} - e^{|t(0) - h(0)|} \quad , \tag{24}$$

and by *(22.a)* one has

$$|1 - a(0)| = |1 - e^{h(0) - t(0)}| \leq e^{|t(0) - h(0)|} - 1 \quad . \tag{25}$$

Combining the inequalities *(21)*, *(24)* and *(25)* proves the theorem.

Q.E.D.

From this theorem it is clear that only a choice of the sequence $\{h(x) : x = 0,1,\ldots\}$ that is close enough to the exact values $\{t(x) : x = 0,1,\ldots\}$, and thus gives rise to a small value of ε, is of practical interest.

In most applications one is not interested in the value of a single probability $p(s)$, but in the cumulative probability $\Pr[S \leq s]$ with $s = 0,1,\ldots$. Note that to simplify the notation only integer s are considered, but a generalization to real s is straightforward. The value of

$\Pr[S \leq s]$ will be calculated as $\displaystyle\sum_{x=0}^{s} f(x)$ and an error bound for the difference between the exact and approximated value is given in the following corollary.

COROLLARY 1. If *(16)* holds, then

$$\left| \Pr[S \leq s] - \sum_{x=0}^{s} f(x) \right| \leq (e^{\varepsilon} - 1) \Pr[S \leq s] \leq e^{\varepsilon} - 1 \qquad s = 0,1,\dots \ . \tag{26}$$

Proof. For $s = 0$, *(26)* follows immediately from *(20)* and *(25)*.

In case s is a strictly positive integer, one gets from *(20)*

$$\Pr[S \leq s] - \sum_{x=0}^{s} f(x) = [1 - a(0)] \Pr[S \leq s] - \sum_{y=1}^{s} \sum_{x=1}^{y} a(x) p(y-x)$$

and thus

$$\left| \text{Prob}[S \leq s] - \sum_{x=0}^{s} f(x) \right| \leq \left[|1 - a(0)| + \sum_{x=1}^{\infty} |a(x)| \right] \Pr[S \leq s]$$

from which by use of *(24)* and *(25)* the inequality *(26)* follows.

Q.E.D.

Remark that the first upper bound given in *(26)* is not computable since $\Pr[S \leq s]$ is assumed to be unknown. However, this bound can be transformed into a computable bound as will be shown in Corollary 2.

COROLLARY 2. If *(16)* holds with $\varepsilon < \ln 2$, then

$$\left| \Pr[S \leq s] - \sum_{x=0}^{s} f(x) \right| \leq \frac{e^{\varepsilon} - 1}{2 - e^{\varepsilon}} \left| \sum_{x=0}^{s} f(x) \right| \qquad s = 0,1,\dots \ . \tag{27}$$

Proof. By *(26)* one has

$$\left| \Pr[S \leq s] - \sum_{x=0}^{s} f(x) \right| \leq (e^{\varepsilon} - 1) \left| \Pr[S \leq s] - \sum_{x=0}^{s} f(x) \right| + (e^{\varepsilon} - 1) \sum_{x=0}^{s} f(x)$$

from which *(27)* follows.

To conclude this section consider the probability $\Pr[S \in J]$, with J a subset of the non-negative integers. This probability will be calculated as $\sum_{s \in J} f(s)$ and an error bound for the approximation is derived in the following corollary.

COROLLARY 3. If *(16)* holds, then

$$\sup_{J} \left| \Pr[S \in J] - \sum_{s \in J} f(s) \right| \le \frac{1}{2} \left[e^{\varepsilon} - 1 + |1 - F(1)| \right] \le e^{\varepsilon} - 1 \quad . \tag{28}$$

Proof. From Theorem 1, one immediately finds that

$$F(1) \le \sum_{s=0}^{\infty} |p(s) - f(s)| + \sum_{s=0}^{\infty} p(s) \le e^{\varepsilon}$$

so that one has for any J

$$2 \left| \Pr[S \in J] - \sum_{s \in J} f(s) \right|$$

$$= \left| \Pr[S \in J] - \sum_{s \in J} f(s) \right| + \left| 1 - \Pr[S \notin J] - F(1) + \sum_{s \in J} f(s) \right|$$

$$\le \left| \Pr[S \in J] - \sum_{s \in J} f(s) \right| + \left| \Pr[S \notin J] - \sum_{s \in J} f(s) \right| + |1 - F(1)|$$

$$\le \sum_{s=0}^{\infty} |p(s) - f(s)| + |1 - F(1)|$$

from which the first inequality follows by application of Theorem 1.

Now, the second upper bound follows from the fact that

$$|1 - F(1)| = \left| \sum_{s=0}^{\infty} [p(s) - f(s)] \right| \le \sum_{s=0}^{\infty} |p(s) - f(s)| \le e^{\varepsilon} - 1 \quad .$$

Q.E.D.

3.   Error Bounds for Stop-Loss Premiums

The net stop-loss premium with retention t of the aggregate claims S will be denoted by $\Pi(t)$. Without loss of generality, t can be assumed to be a positive integer so that

$$\Pi(t) = E[\max(0, S - t)] = \sum_{s=t+1}^{\infty} (s-t)\,p(s) \tag{29}$$

which can be transformed into

$$\Pi(t) = \sum_{s=0}^{t} (t-s)p(s) + E(S) - t \tag{30}$$

In the sequel it will be assumed that the mean $E(S) = \Pi(0)$ can be computed exactly and directly without knowing the aggregate claims distribution. This is the case in the individual model.

Looking at (30), it seems obvious to approximate $\Pi(t)$ by

$$\Omega_1(t) = \sum_{s=0}^{t} (t-s)f(s) + E(S) - t \tag{31}$$

The approximated stop-loss premium $\Omega_1(t)$ can also be computed recursively by means of the scheme

$$\begin{cases} \Omega_1(0) = E(S) & \hspace{6em} (32.a) \\[2ex] \Omega_1(t+1) = \Omega_1(t) + \sum_{s=0}^{t} f(s) - 1 & t = 0,1,\dots \\[2ex] \hspace{3.5em} = 2\Omega_1(t) - \Omega_1(t-1) + f(t) & t = 1,2,\dots \hspace{2em} (32.b) \end{cases}$$

An error bound for the difference between the exact and the approximated value is given in the following theorem.

THEOREM 2. If (16) holds, then

$$\left| \Pi(t) - \Omega_1(t) \right| \le (e^{\varepsilon} - 1)\left[\Pi(t) + t - E(S) \right] \hspace{4em} t = 0,1,\dots \tag{33}$$

Proof. For $t = 0$ (33) is trivial.

In case t is a strictly positive integer, one gets from (20), (30) and (31)

$$\Pi(t) - \Omega_1(t) = [1 - a(0)] \sum_{s=0}^{t} (t-s)p(s) - \sum_{s=1}^{t} (t-s) \sum_{x=1}^{s} a(x)p(s-x) \quad .$$

Taking absolute values gives

$$|\Pi(t) - \Omega_1(t)| \le |1 - a(0)| \sum_{s=0}^{t} (t-s)p(s) + \sum_{x=1}^{t} |a(x)| \sum_{s=0}^{t} (t-s)p(s)$$

$$\le \left[|1 - a(0)| + \sum_{x=1}^{\infty} |a(x)| \right] [\Pi(t) + t - E(S)]$$

and *(33)* follows from *(24)* and *(25)*

$$Q.E.D.$$

The bound *(33)* is not useful in practice since the probabilities $p(s)$ and hence $\Pi(t)$ are assumed to be unknown. A computable bound is given in the following corolarry.

COROLLARY 4. If *(16)* holds with $\varepsilon < \ln 2$, then

$$|\Pi(t) - \Omega_1(t)| \le \frac{e^\varepsilon - 1}{2 - e^\varepsilon} [\Omega_1(t) + t - E(S)] \qquad\qquad t = 0,1,\dots \qquad (34)$$

Proof. The proof is based on the preceding theorem and is similar to the proof of Corollary 2.

$$Q.E.D.$$

Now, an alternative manner to compute an approximative stop-loss premium will be examined. In view of *(29)*, it is natural to consider the approximation

$$\Omega_2(t) = \sum_{s=t+1}^{\infty} (s-t)f(s) \qquad\qquad t = 0,1,\dots \qquad (35)$$

which can be rewritten as

$$\Omega_2(t) = \sum_{s=0}^{t} (t-s)f(s) + F'(1) - tF(1) \quad . \qquad (36)$$

It is clear that *(35)* and *(36)* are only meaningful if $F(1)$ and $F'(1)$ are finite. This will be the case in the applications considered in Section 4 where easy computable expressions for $F(1)$ and $F'(1)$ will be derived.

The approximation $\Omega_2(t)$ can also be computed recursively by the formula

$$\begin{cases} \Omega_2(0) = F'(1) \\ \Omega_2(t+1) = \Omega_2(t) + \sum_{s=0}^{t} f(s) - F(1) \qquad t = 0,1,\dots \qquad (37.a) \\ \qquad\qquad = 2\Omega_2(t) - \Omega_2(t-1) + f(t) \qquad t = 1,2,\dots \qquad (37.b) \end{cases}$$

Remark that *(31)* and *(36)* give rise to the following relation

$$\Omega_2(t) - \Omega_1(t) = F'(1) - E(S) + t[1 - F(1)] \quad , \qquad (38)$$

so that in general $\Omega_2(t)$ will differ from $\Omega_1(t)$. Both definitions will coincide if the approximation is such that $F(1)=1$ and $F'(1)=E(S)$.

Next, the analogue of Theorem 2 is given for the approximation $\Omega_2(t)$.

THEOREM 3. If *(16)* holds and if there exists a real number δ such that

$$\sum_{x=1}^{\infty} x|t(x) - h(x)| \leq \delta \quad , \qquad (39)$$

then

$$|\Pi(t) - \Omega_2(t)| \leq (e^\epsilon - 1)\Pi(t) + \delta e^\epsilon \qquad\qquad t = 0, 1, \ldots \qquad (40)$$

Proof. From *(20)*, *(29)* and *(35)* one obtains

$$\Pi(t) - \Omega_2(t) \leq [1 - a(0)] \sum_{s=t+1}^{\infty} (s-t)p(s) - \sum_{s=t+1}^{\infty} (s-t)\sum_{x=1}^{s} a(x)p(s-x)$$

so that

$$|\Pi(t) - \Omega_2(t)| \leq |1 - a(0)|\Pi(t) + \sum_{x=1}^{\infty} |a(x)| \sum_{s=\max(t+1-x,0)}^{\infty} (s+x-t)p(s)$$

$$\leq |1 - a(0)|\Pi(t) + \sum_{x=1}^{\infty} |a(x)| \sum_{s=t+1}^{\infty} (s-t)p(s) + \sum_{x=1}^{\infty} x|a(x)|$$

$$= \left[|1 - a(0)| + \sum_{x=1}^{\infty} |a(x)|\right]\Pi(t) + \sum_{x=1}^{\infty} x|a(x)|$$

Now, from *(23)* one has that

$$\sum_{x=1}^{\infty} x|a(x)| \leq \exp\left[\sum_{y=0}^{\infty} |t(y) - h(y)|\right] \sum_{x=1}^{\infty} x\, b(x)$$

$$\leq e^\epsilon \sum_{x=1}^{\infty} x|t(x) - h(x)| \leq \delta\, e^\epsilon \qquad (41)$$

which together with *(24)* and *(25)* proves the theorem.

$$Q.E.D.$$

The bound *(40)* can again be transformed into a computable bound. The result is given in the following corollary.

COROLLARY 5. If *(16)* and *(39)* hold with $\varepsilon < \ln 2$, then

$$\left|\Pi(t) - \Omega_2(t)\right| \le \frac{(e^\varepsilon - 1)\Omega_2(t) + \delta\, e^\varepsilon}{2 - e^\varepsilon} \tag{42}$$

Further, remark that since $\Pi(t) \le E(S)$ for every retention t, also the following bound is obtained from Theorem 3.

COROLLARY 6. If *(16)* and *(39)* hold, then

$$\sup_t \left|\Pi(t) - \Omega_2(t)\right| \le (e^\varepsilon - 1)\, E(S) + \delta\, e^\varepsilon \tag{43}$$

Now, the question arises which of the approximations $\Omega_1(t)$ or $\Omega_2(t)$ should be used to approximate $\Pi(t)$. To answer this question, consider the bounds $f_1(t)$ and $f_2(t)$ for $\Omega_1(t)$ and $\Omega_2(t)$ respectively, as given in the Theorems 2 and 3.

$$f_1(t) = (e^\varepsilon - 1)\left[\Pi(t) + t - E(S)\right]$$
$$f_2(t) = (e^\varepsilon - 1)\,\Pi(t) + \delta\, e^\varepsilon$$

Clearly $f_1(t)$ is an increasing function of t, while $f_2(t)$ is a decreasing function of t. It is easily seen that $f_1(t)$ is smaller than $f_2(t)$ if and only if $t < t^*$ with

$$t^* = E(S) + \frac{\delta}{1 - e^{-\varepsilon}} \tag{44}$$

So one finds that for $t < t^*$ the smallest error bound is obtained by using $\Omega_1(t)$, while for $t > t^*$, $\Omega_2(t)$ gives rise to the smallest bound. In practical applications an approximation method will be chosen with small ε and δ, so that the critical value t^* will be approximately equal to $E(S)$.

To conclude this section, consider the problem of calculating a stop-loss premium in case the reinsurer's liability is limited to a certain amount, say $m - t$ with t the retention. The reinsurance risk premium $\Pi(t, m)$ is then given by

$$\Pi(t, m) = \sum_{s=t+1}^{\infty} \min(s - t, m) p(s) = \Pi(t) - \Pi(t + m) \tag{45}$$

Several approximations for $\Pi(t, m)$ can be conceived. Since in most applications the retention t will be larger than $E(S)$, the approximation $\Omega(t, m)$ of $\Pi(t, m)$ will be defined in terms of $\Omega_2(t)$ and $\Omega_2(t + m)$ by

$$\Omega(t, m) = \sum_{s=t+1}^{\infty} \min(s - t, m)\, f(s) = \Omega_2(t) - \Omega_2(t + m) \tag{46}$$

Theorem 3 can be generalized, as is shown in Theorem 4.

THEOREM 4. If *(16)* and (39) hold, then

$$|\Pi(t,m) - \Omega(t,m)| \leq (e^{\varepsilon} - 1)\, \Pi(t,m) + \delta\, e^{\varepsilon} \qquad (47)$$

Proof. Proceeding as in the proof of Theorem 3, one obtains

$$|\Pi(t,m) - \Omega(t,m)| \leq |1 - a(0)|\, \Pi(t,m) + \sum_{x=1}^{\infty}|a(x)| \sum_{s=\max(t+1-x,0)}^{\infty} \min(s+x-t,m)\, p(s)$$

since $\min(s+x-t,m) \leq x + \min(s-t,m)$, one has

$$|\Pi(t,m) - \Omega(t,m)| \leq \left[|1 - a(0)| + \sum_{x=1}^{\infty}|a(x)|\right] \Pi(t,m) + \sum_{x=1}^{\infty} x|a(x)|$$

which together with *(24)*, *(25)* and *(41)* proves the theorem.

Q.E.D.

Once again, *(47)* can be transformed into a computable bound, as is shown in the following corollary.

COROLLARY 7. If *(16)* and *(39)* hold with $\varepsilon < \ln 2$, then

$$|\Pi(t,m) - \Omega(t,m)| \leq \frac{(e^{\varepsilon} - 1)\, \Omega(t,m) + \delta\, e^{\varepsilon}}{2 - e^{\varepsilon}} \qquad (48)$$

4. Applications

The preceding general results permit to derive error bounds for the approximation methods for the individual model proposed by KORNYA (1983), HIPP (1986) and DE PRIL (1988 and 1989). For these methods some partial results can already be found in the literature. An overview is given in Table 1.

	KORNYA	HIPP	DE PRIL
THEOREM 1 or a related result	KORNYA (1983) HIPP (1986) DE PRIL (1988 & 1989)	HIPP (1986) DE PRIL (1988 & 1989) HIPP & MICHEL (1990)	DE PRIL (1988 & 1989)
THEOREM 3	HIPP (1986)	—	—

Table 1. *Overview of the Literature*

The following remarks can be made about the given references :
  i) The first bound for the difference between the exact probabilities and their approximated values has been obtained by KORNYA (1983) _ and the discussants of that paper _ but this bound can be improved, see DE PRIL (1988 and 1989).
  ii) KORNYA (1983) and DE PRIL (1988) only consider the individual life model, while the other papers deal with the general case of an individual model with arbitrary claims.
  iii) The bounds for the Kornya approximation in HIPP (1986, formulas *(5)* and *(10)*) contain an error. The correct bounds follow by application of the Theorems 1 and 3 and the forthcoming Table 2. See also DE PRIL (1988 and 1989).
  iv) HIPP (1986) and HIPP and MICHEL (1990) also present some bounds in terms of concentration functions. Although interesting from a mathematical point of view, these bounds will not be discussed here since they are hard to compute in practical applications. See DE PRIL and DHAENE (1992) for further considerations.

The objective of the present paper is to treat the different approximation methods in a general and unified way. Bounds for the different methods follow immediately from one single theorem. The new results mainly concern the error bounds for the stop-loss premiums (Theorems 2, 3 and 4).

Now, the different approximations will be discussed in more detail. For convenience they will be treated in reversed chronological order.

Besides the notations introduced in Example 2, the symbol μ_i will be used to denote the conditional expected claim amount of policy i, given that a claim occurs.

$$\mu_i = \sum_{x=1}^{\infty} x\, g_i(x) \qquad\qquad i = 1, 2, \dots, n \qquad (49)$$

Further, it will be assumed that for each policy i the claim probability q_i is smaller than ½ .

4.1.   The Approximations of DE PRIL

From the assumption $q_i < 1/2$, $i = 1, 2, \dots, n$, it follows that $(q_i/p_i)^k$ becomes smaller and smaller for increasing k. A natural way of approximating the coefficients $t(x)$ is to restrict the summation over k in *(10.b)* to a summation over all k-values smaller than or equal to r, with r some positive integer. Given r, the $t(x)$ are thus approximated by

$$
\begin{cases}
h^{(r)}(0) = \displaystyle\sum_{i=1}^{n} \ln p_i & (50.a) \\[2mm]
h^{(r)}(x) = \displaystyle\sum_{k=1}^{\min(r,x)} \frac{(-1)^{k+1}}{k} \sum_{i=1}^{n} \left(\frac{q_i}{p_i}\right)^k g_i^{*k}(x) & x = 1, 2, \dots \quad (50.b)
\end{cases}
$$

In *(50)* a superscript (r) is added to the symbol $h(x)$ in order to specify more explicitly the approximation under consideration. For the same reason, the notations $f(s)$, $\Omega_1(t)$, $\Omega_2(t)$, $\Omega(t,m)$, ε, δ, $F(1)$ and $F'(1)$ used in the preceding sections, will be altered in $f^{(r)}(s)$, $\Omega_1^{(r)}(t)$, $\Omega_2^{(r)}(t)$, $\Omega^{(r)}(t,m)$, $\varepsilon(r)$, $\delta(r)$, $F^{(r)}(1)$ and $F'^{(r)}(1)$ respectively. In the sequel r will be called the order of the approximation.

Inserting the $h^{(r)}(x)$ in *(12)* leads to a recursive scheme for computing the *r-th* order approximations $f^{(r)}(s)$. It is immediately clear that $f^{(r)}(s) = p(s)$ for $s \le r$.

The key quantities needed to compute the error bounds derived in the preceding sections are summarized in Table 2. The expressions for $\varepsilon(r)$ and $\delta(r)$ follow immediately from *(10)* and *(50)*. Expression for $F^{(r)}(1)$ and $F'^{(r)}(1)$ can be obtained from the following lemma (in which the index (r) has been omitted).

LEMMA 1. If the sequence $\{h(x): x = 0,1,...\}$ fulfils the condition

$$\sum_{x=1}^{\infty} x\,|h(x)| < \infty \tag{51}$$

then the series $\sum_{s=0}^{\infty} f(s)$ and $\sum_{s=1}^{\infty} s\, f(s)$ are absolute convergent and have as sum respectively

$$F(1) \;=\; \exp\left[\sum_{x=0}^{\infty} h(x)\right] \tag{52}$$

and

$$F'(1) \;=\; F(1)\sum_{x=1}^{\infty} x\, h(x) \tag{53}$$

Proof. From *(7)* and *(51)* it follows that the quantities $\{d(s): s = 0,1,...\}$ defined by the recursion

$$\begin{cases} d(0) \;=\; \exp\left[-\sum_{x=1}^{\infty} |h(x)|\right] \\[2mm] s\, d(s) \;=\; \sum_{x=1}^{s} x\,|h(x)|\, d(s-x) \end{cases} \qquad s = 1,2,...$$

can be considered as probabilities associated with a compound Poisson distributed random variable with finite mean.

Using *(12)* it is easily shown by induction that

$$|f(s)| \;\le\; d(s)\exp\left[\sum_{x=0}^{\infty} |h(x)|\right]$$

so that $\sum_{s=0}^{\infty} f(s)$ and $\sum_{s=1}^{\infty} s\, f(s)$ converge absolutely.

Now, the expressions *(52)* and *(53)* follow from *(14)*, *(15)* and Abel's limit theorem.

Q.E.D.

It is easy to verify that in case the $h^{(r)}(x)$ are defined by *(50)*, the condition *(51)* is satisfied. The corresponding values *(52)* and *(53)* for $F^{(r)}(1)$ and $F'^{(r)}(1)$ are given in Table 2.

4.2.   The Approximations of HIPP

In order to define Hipp's approximations write $\ln P(u)$ as

$$\ln P(u) = \sum_{i=1}^{n} \ln \left[1 + q_i (G_i(u) - 1)\right]$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} q_i^k (G_i(u) - 1)^k$$

This expansion is of the form *(2)* with coefficients $t(x)$ given by

$$\begin{cases} t(0) = -\sum_{k=1}^{\infty} \sum_{i=1}^{n} \frac{q_i^k}{k} & (54.a) \\[4mm] t(x) = \sum_{k=1}^{x} \sum_{l=1}^{k} \frac{(-1)^{l+1}}{k} \binom{k}{l} \sum_{i=1}^{n} q_i^k g_i^{*l}(x) & x = 1, 2, \ldots & (54.b) \end{cases}$$

Now, the approximations $h^{(r)}(x)$ are defined by neglecting in *(54)* the terms in q_i^k for $k > r$.

$$h^{(r)}(0) = -\sum_{k=1}^{r} \sum_{i=1}^{n} \frac{q_i^k}{k} \qquad\qquad (55.a)$$

$$h^{(r)}(x) = \sum_{k=1}^{\min(r,x)} \sum_{l=1}^{k} \frac{(-1)^{l+1}}{k} \sum_{i=1}^{n} q_i^k g_i^{*l}(x) \qquad x = 1, 2, \ldots \qquad (55.b)$$

Inserting these coefficients in *(12)* leads to a recursive scheme for computing the *r-th* order approximations suggested by HIPP (1986).

Error bounds can be derived by using the results of the previous sections and the values of $\varepsilon(r)$, $\delta(r)$, $F^{(r)}(1)$ and $F'^{(r)}(1)$ given in Table 2. Finally, remark that one gets from *(38)* and Table 2 that $\Omega_1^{(r)}(t) = \Omega_2^{(r)}(t)$ and that $\lim_{r \to \infty} f^{(r)}(s) = p(s)$.

4.3.   The Approximations of KORNYA

In order to describe Kornya's method, remark that the coefficients $t(x)$ defined in *(10)* can also be written as

$$
\begin{cases}
t(0) = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \sum_{i=1}^{n} \left(\frac{q_i}{k} \right)^k & (56.a) \\
\\
t(x) = \sum_{k=1}^{x} \frac{(-1)^{k+1}}{k} \sum_{i=1}^{n} \left(\frac{q_i}{p_i} \right)^k g_i^{*k}(x) \qquad x = 1,2,\ldots & (56.b)
\end{cases}
$$

These coefficients are approximated by coefficients $h^{(r)}(x)$, obtained by neglecting in *(56)* the terms in $\left(\dfrac{q_i}{p_i} \right)^k$ for $k > r$.

$$
\begin{cases}
h^{(r)}(0) = \sum_{k=1}^{r} \frac{(-1)^k}{k} \sum_{i=1}^{n} \left(\frac{q_i}{k} \right)^k & (57.a) \\
\\
h^{(r)}(x) = \sum_{k=1}^{\min(r,x)} \frac{(-1)^{k+1}}{k} \sum_{i=1}^{n} \left(\frac{q_i}{p_i} \right)^k g_i^{*k}(x) \qquad x = 1,2,\ldots & (57.b)
\end{cases}
$$

The related values of $\varepsilon(r)$, $\delta(r)$, $F^{(r)}(1)$ and $F^{\prime(r)}(1)$ are again given in Table 2. Remark also that the approximations are asymptotically correct : $\lim_{r \to \infty} f^{(r)}(s) = p(s)$.

	KORNYA	HIPP	DE PRIL
$\varepsilon\,(r)$	$\dfrac{1}{r+1}\sum_{i=1}^{n}\left(p_i+\dfrac{p_i}{p_i-q_i}\right)\left(\dfrac{q_i}{p_i}\right)^{r+1}$	$\dfrac{1}{r+1}\sum_{i=1}^{n}\dfrac{(2q_i)^{r+1}}{p_i-q_i}$	$\dfrac{1}{r+1}\sum_{i=1}^{n}\dfrac{p_i}{p_i-q_i}\left(\dfrac{q_i}{p_i}\right)^{r+1}$
$\delta\,(r)$	$\sum_{i=1}^{n}\mu_i\dfrac{p_i}{p_i-q_i}\left(\dfrac{q_i}{p_i}\right)^{r+1}$	$\dfrac{1}{2}\sum_{i=1}^{n}\mu_i\dfrac{(2q_i)^{r+1}}{p_i-q_i}$	$\sum_{i=1}^{n}\mu_i\dfrac{p_i}{p_i-q_i}\left(\dfrac{q_i}{p_i}\right)^{r+1}$
$F^{(r)}(1)$	1	1	$p(0)\exp\left[\sum_{k=1}^{r}\dfrac{(-1)^{k+1}}{k}\sum_{i=1}^{n}\left(\dfrac{q_i}{p_i}\right)^{k}\right]$
$F'^{(r)}(1)$	$\sum_{i=1}^{n}\mu_i\left[q_i+p_i\left(-\dfrac{q_i}{p_i}\right)^{r+1}\right]$	$E\,(S)$	$F^{(r)}(1)\sum_{i=1}^{n}\mu_i\left[q_i+p_i\left(-\dfrac{q_i}{p_i}\right)^{r+1}\right]$

Table 2. Key Values of Some Approximation Methods

4.4.   Comparison of the Different Methods

The following observations can be made about the approximation methods under consideration.

   i)   The three methods lead to asymptotically correct approximations. Moreover in De Pril's case the *r-th* order approximations are exact for the first *r* values.

   ii)   Table 2 indicates that the approximations $f^{(r)}(s)$ of Kornya and Hipp are normed to 1 and that Hipp's approximations are the only ones that preserve the mean. However, remark that for each of the methods the approximations $f^{(r)}(s)$ can be negative for some values of s , so that these properties seem not to be so important after all.

   iii)   From *(50)* and *(57)* it follows that the methods of Kornya and De Pril only differ by the choice of the starting value $h^{(r)}(0)$.

   iv)   The recursive formulas of Kornya and De Pril require the same amount of computing time. Comparison of *(55)* with *(50)* or *(57)* shows that Hipp's method necessitates more computational effort since the $h^{(r)}(x)$ contain an additional summation.

   v)   The magnitude of the error bounds is determined by the values of $\varepsilon\,(r)$ and $\delta\,(r)$. From Table 2 it is clear that

$$\text{for } \varepsilon\,(r): \text{ De Pril} < \text{Kornya} < \text{Hipp};$$
$$\text{for } \delta\,(r): \text{ De Pril} = \text{Kornya} < \text{Hipp}.$$

By making a choice between these methods one should compare their accuracy and the computation effort they require. From the above comparison it follows that the approximations of De Pril minimize the theoretical error bounds and the computation effort at the same time. Therefore, this method seems to be the most suited for the practical computation of the aggregate claims distribution and stop-loss premiums.

4.5.   Compound Poisson Approximations

Let λ_i with $i = 1,2,...,n$, denote positive real numbers and put $\lambda = \sum_{i=1}^{n} \lambda_i$. Consider the sequence $\{h(x): x = 0,1,...\}$ defined as

$$
\begin{cases}
h(0) = -\lambda & \quad (58.a) \\[2mm]
h(x) = \sum_{i=1}^{n} \lambda_i\, g_i(x) & \qquad x = 1,2,... \quad (58.b)
\end{cases}
$$

Insertion of these coefficients in (12) leads to the following approximations $\{f(s): s = 0,1,...\}$

$$f(0) = e^{-\lambda} \tag{59.a}$$

$$s\,f(s) = \sum_{x=1}^{s} x \sum_{i=1}^{n} \lambda_i g_i(x) f(s-x) \qquad s = 1,2,... \tag{59.b}$$

From (7) it is clear that $f(s)$ is the probability function of a compound Poisson distributed variable with Poisson parameter λ and claim amount distribution

$$g(x) = \frac{1}{\lambda} \sum_{i=1}^{n} \lambda_i g_i(x) \tag{60}$$

In this way, each choice of the parameters λ_i leads to a compound Poisson approximation of the individual risk model.

Remark that the choice $\lambda_i = q_i / p_i$ corresponds to the first order approximation of KORNYA (1983) and that the most common choice $\lambda_i = q_i$ gives rise to the first order approximation of HIPP (1986).

Error bounds for the general class of compound Poisson approximations defined by (59) are considered in DE PRIL and DHAENE (1992). A comparison of these results with the ones obtained in the present paper for the first order approximations of Kornya and Hipp reveals that the bounds given in DE PRIL and DHAENE (1992) are sharper. 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 here. If only a crude first order approximation for the distribution function or stop-loss premiums is required, one should refer to DE PRIL and DHAENE (1992).

5. References

DE PRIL, N. (1986). On the exact computation of the aggregate claims distribution in the individual life model. *ASTIN Bulletin,* 16(2), 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(1), 9-24.

DE PRIL, N. and DHAENE, J. (1992). Error bounds for compound Poisson approximations of the individual risk model. *ASTIN Bulletin,* 22(2), 135-148.

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

HIPP, C. and MICHEL, R. (1990). *Risicotheorie : Stochastische Modelle und Statistische Methoden.* Schriftenreihe Angewandte Versicherungsmathematik, Heft 24, Deutsche Gesellschaft für Versicherungsmathematik, Verlag Versicherungswirtschaft e.V., Karlsruhe.

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.

KUON, S., REICH, A. and REIMERS, L. (1987). Panjer vs. Kornya vs. De Pril : a comparison from a practical point of view. *ASTIN Bulletin,* 17(2), 183-191.

PANJER, H.H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin,* 12(1), 22-26.