DISCUSSION PAPERS

MAXIMIZING COMPOUND POISSON STOP-LOSS PREMIUMS NUMERICALLY WITH GIVEN MEAN AND VARIANCE

BY R. KAAS
University of Amsterdam

M. VANNESTE [1]
K.U. Leuven

AND M.J. GOOVAERTS
K.U. Leuven, University of Amsterdam

ABSTRACT

This paper describes a technique to find the maximal stop-loss premiums in a given retention for a compound Poisson risk with known parameter, and known mean and variance of the claims. Restricting to an arithmetic and finite support of the claims, one gets an optimization problem of a non-linear function with a computable gradient, under linear constraints.

Numeraical results are given contrasting the method with the method of a previous paper, where only diatomic distributions were considered.

1. INTRODUCTION

At the XXII ASTIN-colloquium, Hans Schmitter provoked a discussion on the problem of maximizing the ruin probability with given initial capital d and safety margin θ, and for the individual claims mean μ, variance σ^2 and known maximum b. In the article by KAAS (1991) this problem is tackled by restricting the feasible claim amount distributions to be diatomic. Under these circumstances one may use one-dimensional optimization techniques to find the optimal diatomic distribution. It is stated in that paper that the maximizing diatomic claim distributions tend to be so good that if one tries many random feasible solutions, only occasionally a higher ruin probability is found. The same restriction to diatomic distributions also leads to good results for the closely related problem of finding maximal compound Poisson stop-loss premiums. In this paper we discuss a more sophisticated technique, only usable for the latter problem, of finding the maximal stop-loss premium without the restriction to diatomic distributions. The results indicate, however, that this restriction is not very severe, so practically oriented readers do wise to use the much simpler diatomic method of KAAS (1991) instead of the one discussed here.

In the case of an n-point support for the claims distribution, there are n mass points and n probabilities, but three restrictions, viz. the sum of the probabil-

[1] Acknowledgement: M. Vanneste would like to thank the support of grant OT/90/2 of K.U.Leuven.

ities equals one, the mean claim is μ, and the mean squared claim is $\mu^2+\sigma^2$. So we are left with $2n-3$ free variables, over which the compound Poisson stop-loss premium is to be maximized. Now suppose we assume the support to be fixed in advance. Then the resulting restrictions on the probabilities are all linear. We will capitalize on this fact in choosing an appropriate solution technique from the field of operations research. If moreover we choose the support to be arithmetic, we can exploit Panjer's recursion algorithm to find the compound Poisson probabilities effortlessly. No such recursion exists to compute the exact ruin probability, so this method cannot be used for Schmitter's problem.

In Section 2 we present the mathematical model. In Section 3 we shortly describe the optimization technique used, and show how our problem, to be given in (2.7)-(2.11), can be rewritten in a suitable form, with as few variables and restrictions as possible. In Section 4 we give some examples for small n, which contributes to the insight in why this problem is a tough one. Note that since the objective is to approximate the unrestricted optimum, the method is only meaningful for a fine discretization, which means a large n. In Section 5 we expand on the numerical examples of KAAS (1991), and try to determine whether his diatomic optima are far removed from the global optima.

2. THE MATHEMATICAL MODEL

We study the following random variable, representing the total claims:

$$S = X_1+X_2+ \ldots +X_N \tag{2.1}$$

where the random variable N describes the number of claims and is Poisson (λ) distributed. The random variables X_i are *iid* and describe the claim sizes. They are assumed independent of N. Their distribution is arithmetic with span h and maximum $b = n\cdot h$, so $P[X_i \in \{0, h, 2h, \ldots, nh\}] = 1$. In this note we will not study the role of the maximum claim size closely, which means we take b to be 'large enough'. We assume that $E[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2$, where μ and σ^2 are known.

We use the following notation for the probability functions of X_i and S:

$$p_j = P[X_i = jh] \qquad j = 0, 1, \ldots, n \tag{2.2}$$

$$f_j = P[S = jh] \qquad j = 0, 1, \ldots \tag{2.3}$$

The probabilities f_j can be computed from the p_j's by Panjer's recursion:

$$f_0 = e^{\lambda(p_0-1)} \tag{2.4}$$

$$f_j = \frac{\lambda}{j}\sum_{i=1}^{j} i\cdot p_i\cdot f_{j-i} \qquad j = 1, 2, \ldots \tag{2.5}$$

The stop-loss premium of S can be computed as follows:

$$\pi_S(d) = \sum_{s>d/h} (h\cdot s-d)\cdot f_s = E[S]-d+\sum_{s=0}^{[d/h]} (d-h\cdot s)\cdot f_s \tag{2.6}$$

Of course by (2.1) we know that $E[S] = \lambda\mu$. So to find maximal compound Poisson stop-loss premiums $\pi_S(d)$ we only have to maximize the last sum of (2.6) over the feasible values of p_j. Writing $g(\cdot)$ for this function, we must solve

$$\text{(2.7)} \qquad \text{Max } g(p_0, \ldots, p_n) = \text{Max} \sum_{s=0}^{[d/h]} (d - h \cdot s) \cdot f_s$$

with f_s determined by (2.4)-(2.5). The conditions on the probabilities p_j, $j = 0, \ldots, n$, namely $\Sigma p_j = 1$, $E[X] = \mu$ and $\text{Var}[X] = \sigma^2$, can be written down as follows for a claims distribution on mass points $0, h, \ldots, n \cdot h$:

$$\text{(2.8)} \qquad p_0 + \quad p_1 + \ldots + \quad p_n = 1$$

$$\text{(2.9)} \qquad 0 \cdot p_0 + 1 \cdot p_1 + \ldots + n \cdot p_n = \mu/h$$

$$\text{(2.10)} \qquad 0^2 \cdot p_0 + 1^2 \cdot p_1 + \ldots + n^2 \cdot p_n = (\mu^2 + \sigma^2)/h^2$$

Obviously we must require

$$\text{(2.11)} \qquad p_j \geq 0, \qquad j = 0, \ldots, n$$

Note that it is easy to incorporate information on higher moments than the second, though this is not very important for practical purposes.

3. THE METHOD OF THE LINEAR COMBINATIONS

To solve the problem presented in the previous section we use a so-called gradient method. In these methods one generates a sequence of feasible solutions, choosing the next point in a direction determined with the help of the gradient of the objective function. This is done in such a way that the objective function increases. The iteration is stopped if such improvement is no longer possible. The method we use is the method of the linear combinations, see TAHA (1987, § 19.2.5). We give a short description of it, which should be adequate for those with some knowledge of operations research. A good introduction in this subject is TAHA (1987). See also Figure 1.

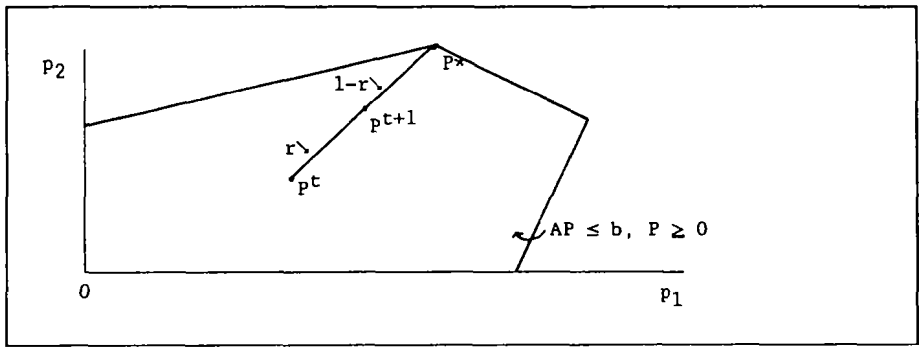

FIGURE 1. The method of linear combinations for $n = 2$.

Suppose the vector of variables is $P = (p_0, p_1, \ldots, p_n)'$. We want to determine the maximum of $g(P)$:

$$\text{Max } g(P) \tag{3.1}$$

under the following linear restrictions on P:

$$\begin{aligned} AP &\leq b \\ P &\geq 0 \end{aligned} \tag{3.2}$$

Starting from a feasible solution P^0, in the $t+1$st iteration of this method P^{t+1} is determined as a linear combination of two feasible solutions:

$$P^{t+1} = P^t + r(P^* - P^t) \tag{3.3}$$

for some r with $0 < r \leq 1$, with P^* the vertex of the feasible region which gives the optimal solution to the following linear programming problem:

$$\max_P \ \nabla g(P^t) \cdot P \tag{3.4}$$

under the same conditions (3.2). The value of r in (3.3) is taken to be the optimal solution to the one-dimensional optimization problem

$$\max_{r \in (0,\,1]} \ h(r) = g(P^t + r(P^* - P^t)) \tag{3.5}$$

so r is the 'best' step-size in (0, 1], if a step is taken in the direction of P^*. Because the feasible region is convex, point P^{t+1} is also a feasible solution. This procedure is repeated until

$$\nabla g(P^t) \cdot P^* \leq \nabla g(P^t) \cdot P^t \tag{3.6}$$

In this point improvement can no longer be found by this method. The point P^t might, however, well be a local maximum, so it is generally wise to repeat the whole process several times with different starting solutions P^0.

We can get rid of the three linear equalities in problem (2.6)-(2.11) by elimination, and write it in the form (3.1)-(3.2). As dependent variables we choose p_0, the probability of a zero-claim, p_n, the probability of the maximum value $b = h \cdot n$ and p_m, the probability of some point near the mean claim: $h \cdot m \approx \mu$. This choice has the advantage that in general a feasible solution is found by taking all other probabilities equal to zero.

Write $V = \{0, 1, \ldots, n\} \backslash \{0, m, n\}$ for the set indices of free variables, then (2.11) first gives the $n-2$ non-negativity conditions

$$p_i \geq 0 \qquad \text{for} \qquad i \in V; \tag{3.7}$$

the constraints $p_n \geq 0$, $p_m \geq 0$ and $p_0 \geq 0$ directly lead to three linear inequal-

ities in the probabilities p_i, $i \in V$, namely

$$(3.8)\quad p_n = \frac{1}{n^2 - nm}\left\{(\sigma^2+\mu^2)/h^2 - m\mu/h - \sum_{i\in V} i^2 p_i + m\sum_{i\in V} i p_i\right\} \geq 0$$

$$(3.9)\quad p_m = \frac{1}{m^2 - nm}\left\{(\sigma^2+\mu^2)/h^2 - n\mu/h - \sum_{i\in V} i^2 p_i + n\sum_{i\in V} i p_i\right\} \geq 0$$

$$(3.10)\quad p_0 = 1 - \sum_{i\in V} p_i - p_n - p_m \geq 0$$

where obviously (3.8) and (3.9) have to be inserted in (3.10). These three inequalities are the constraints of the linear problem (3.4) that must be solved in each iteration. The objective function is

$$(3.11)\quad w_t(p_1, \ldots, p_{m-1}, p_{m+1}, \ldots, p_{n-1}) := (\nabla g)\cdot(p_1, \ldots, p_{m-1}, p_{m+1}, \ldots, p_{n-1})'$$

The gradient ∇g contains the derivatives of $g(\cdot)$ in (2.7) with respect to the free variables p_k, $k \in V$. The necessary partial derivatives of f_s can be determined from the recursive relations (2.5):

$$(3.12)\quad \frac{\partial f_0}{\partial p_k} = e^{\lambda(p_0-1)}\,\lambda\,\frac{\partial p_0}{\partial p_k}$$

$$\frac{\partial f_j}{\partial p_k} = \lambda/j \sum_{i=1}^{j} i\left[p_i \frac{\partial f_{j-i}}{\partial p_k} + \frac{\partial p_i}{\partial p_k} f_{j-i}\right], \qquad j = 1, 2, \ldots$$

with $\partial p_i/\partial p_k$ equal to the coefficient of p_k in (3.8)-(3.9) for $i \notin V$, equal to one for $i = k$, and zero otherwise.

Note that if (3.12) is computed successively for each $k \in V$, $k \cdot h \leq d$, it suffices to store a vector $\partial f_i/\partial p_k$, $i = 0, \ldots, [d/h]$, rather than a matrix. Furthermore it looks as if the number of computations involved in (3.12) increases with the third power of the number of mass points, so with h^{-3}. Most of the terms in (3.12), however, can be shown to be equal to zero. The first term in the summation only contributes if $p_i > 0$, which is only for the mass points of the current solution P^t. There are three of these if P^t is a vertex of the admissible region, six or less if it is a combination of two vertices, and so on. The partial derivative in the second term of the summation is only non-zero for $i \in \{n, m, k\}$.

The other maximization, required in each iteration step to compute the optimum step-size, is the one-dimensional maximization of (3.5). This can easily be performed using for instance golden section search or Brent's method, see PRESS et al. (1986, Ch. 10.1).

4. ANALYSIS OF THE PROBLEM FOR SMALL VALUES OF n

As stated before the method is meaningful for small values of h only, which indicates that the number of mass points n should be large. But for $n = 3$ and

d small, we can write down explicit analytical expressions for the compound Poisson stop-loss premium. There remains only one free variable, for instance p_0, over which to optimize. The feasible region is a closed interval. We examined the case $\mu = 2, \sigma^2 = 1, \lambda = 1, h = 1$ (so $b = 3$, and the range of the claims is $\{0, 1, 2, 3\}$) and various integer values of d. To determine which values of p_j are feasible, we first solve (3.8)-(3.10) for $p_0 = 0$, and get $p_1 = \frac{1}{2}$, $p_2 = 0$ and $p_3 = \frac{1}{2}$. For $p_1 = 0$ we obtain $p_0 = \frac{1}{6}$, $p_2 = \frac{1}{2}$ and $p_3 = \frac{1}{3}$. All admissible solutions are convex combinations of these two solutions. So the feasible region is $p_0 \in [0, \frac{1}{6}]$, and we have $p_1 = \frac{1}{2} - 3p_0$, $p_2 = 3p_0$ and $p_3 = \frac{1}{2} - p_0$. With (2.4) and (2.5) we can express $g(\cdot)$, see (2.7) and (3.1), in p_0. This leads to:

$d = g(p_0) =$	Maximum for:
1: $e^{p_0 - 1}$	$p_0 = \frac{1}{6}$
2: $e^{p_0 - 1}(2\frac{1}{2} - 3p_0)$	$p_0 = 0$: $g' \leq 0$ for $p_0 \in [0, \frac{1}{6}]$
3: $e^{p_0 - 1}(4\frac{1}{8} - 4\frac{1}{2}p_0 + 4\frac{1}{2}p_0^2)$	$p_0 = \frac{1}{6}$; $g'(0) < 0$, $g'(\frac{1}{6}) > 0$
4: $e^{p_0 - 1}(6\frac{3}{8} - 5\frac{7}{8}p_0 + 2\frac{1}{4}p_0^2 - 4\frac{1}{2}p_0^3)$	Interior maximum: $g'(0) < 0$, $g'(\frac{1}{6}) > 0$
7, 10, 13:	Also interior maxima
5, 6, 8, 9, 11, 12, 14, 15, ...	Boundary maximum $p_0 = 0$

For each integer d, (2.7) is a product of $e^{p_0 - 1}$ times a polynominal in p_0 of degree $d-1$. For non-integer d the values of (2.7) as given above can be found by linear interpolation. For different values of d, different optima arise, following a rather erratic pattern. For large d, one expects $p_0 = 0$ to be the optimal solution, since it maximizes the skewness of S, see KAAS (1991) and GOOVAERTS et al. (1990).

To optimize over p_0 proves worthwhile: the relative difference between minimal and maximal stop-loss premium for $d = 3$, 4, and 5 was 10%, 33% and 21% respectively.

Taking $n = 4$, we obtain a two-dimensional problem, for instance in p_1 and p_3. Optimal solutions can now be found using graphical methods. The objective function is a product of a polynominal in p_1 and p_3, and the exponent of some linear form in p_1 and p_3. As a rule, maxima for (2.7) will be found in the vertices of the admissible region. A plot of the case $\mu = 3$, $\sigma^2 = 1$, $\lambda = 1$, $d = 10$ and $h = 1$ can be seen in Figure 1.

5. NUMERICAL RESULTS FOR LARGE n

In KAAS (1991) the claims distribution is not restricted to be arithmetic. He considers only the diatomic distributions, which makes it possible to solve the problem using one-dimensional optimization techniques. One would think that applying the method of this paper for large values of n, maximal stop-loss premiums would generally be found that are substantially better than those obtained under the severe restriction to diatomic solutions. This proved to be false. For some cases it proved that the diatomic optima were also global optima, see below. For other cases some improvement could be found, but not a substantial one.

In our computations it emerged that quite often the optimal solutions were vertices of the feasible region (3.7)-(3.10) that is, solutions with only three mass points. But the objective function (2.7) can be very irregular, because the polynomial component has degree $[d/h]-1$, so there might be about $\frac{1}{2}[d/h]-1$ local maxima to consider, apart from all the boundary maxima.

Sometimes two of the three mass points appeared to be neighbors in the set of feasible mass points $k \cdot h$, $k = 0, 1, \ldots$. This remained so when the span h was refined, and then these mass points collapsed into one intermediate mass point. Thus a two-point solution proved to be optimal. We consider the examples of KAAS (1991) more closely. His optimal diatomic solutions for the case $\mu = 3$, $\sigma^2 = 1$, $\lambda =$ are given below:

$d = 2$		$d = 7$		$d = 20$	
$\pi_S^* = 4.332192$		$\pi_S^* = 1.395435$		$\pi_S^* = 0.052178$	
$x_1 = 0.4054$	$p_1 = .1293$	$x_1 = 2.6667$	$p_1 = .9000$	$x_1 = 2.9559$	$p_1 = .9981$
$x_2 = 3.3854$	$p_2 = .8707$	$x_2 = 6.0000$	$p_2 = .1000$	$x_2 = 25.674$	$p_2 = .0019$

A better solution was found for the case $d = 7$; it contains a mass point $x_0 = 0$. If the mass on x_0 is p_0, an identical problem arises by replacing the Poisson parameter λ by $\lambda(1-p_0)$, the mean μ by $\mu/(1-p_0)$ and the variance σ^2 by $\sigma^2/(1-p_0)-p_0\{\mu/(1-p_0)\}^2$. By varying p_0 over $p_0 \in [0, \sigma^2/(\mu^2+\sigma^2)]$ one can find the best maximal diatomic solution, having two *positive* mass points. Allowing mass on d or any point larger than d also only requires a minor modification of the method.

In the case at hand we found by trial and error that the best choice is $p_0 = 0.03$. The result:

$d = 7$	
$\pi_S^* = 1.399613$	
$x_0 = 0.0000$	$p_0 = .0300$
$x_1 = 2.7971$	$p_1 = .8680$
$x_2 = 5.6087$	$p_2 = .1020$

Below we give the results for the algorithm from this paper, for the discretization span h we tried that led to the highest maximal stop-loss premium.

$d = 2$, $h = .0204$		$d = 7$, $h = 0.39958$		$d = 20$, $h = .1478$	
$\pi_S^* = 4.332191$		$\pi_S^* = 1.399609$		$\pi_S^* = 0.052178$	
$x_1 = 0.4080$	$p_1 = .1296$	$x_1 = 0.0400$	$p_1 = .0309$	$x_1 = 2.8082$	$p_1 = .0001$
$x_2 = 3.3660$	$p_2 = .0257$	$x_2 = 2.7971$	$p_2 = .8660$	$x_2 = 2.9560$	$p_2 = .9979$
$x_3 = 3.3864$	$p_3 = .8448$	$x_3 = 5.5941$	$p_3 = .1030$	$x_3 = 25.717$	$p_3 = .0019$

It can be seen that the best diatomic solutions for $d = 2$ and $d = 20$ are indeed global optima; the same holds for the solution with $p_0 = 0.03$ at $d = 7$. The best arithmetic claims distribution is very similar to the best diatomic one.

The spans h were optimal in the sense that both a small increase and a small decrease gave worse results. Note that the problem can be solved in theory by

each time halving the value of h. For increasing values of $n = b/h$, however, the computing process proved to become difficult for two reasons. First, the time required to compute the gradient is quadratic to cubic in n. Second, the number of possible local maxima increases with n, also.

For the other numerical example in KAAS (1991), with $\mu = 3$, $\sigma^2 = 1$ and $\lambda = 5$, and $d = 5, 20, 40$, the results were analogous. In this case, too, for the middle retention $d = 20$ it was best to take $p_0 > 0$. In this case $p_0 = 0.036$ led to a maximal stop-loss premium of 1.139811 (for $p_0 = 0$: 1.136463).

An example where a diatomic solution, or a solution with two *positive* mass points, did not prove to be optimal was found taking $\mu = 10$, $\sigma^2 = 3$, $\lambda = 2$ and $d = 60$. The best solution with two positive mass points and the best one found with the method described above are:

Diatomic solution $\pi_S^* = 0.107084$		Arithmetic with $h = .4925$ $\pi_S^* = 0.108495$	
$x_1 = 9.9514$	$p_1 = .9992$	$x_1 = 9.8500$	$p_1 = .9460$
$x_2 = 71.730$	$p_2 = .0008$	$x_2 = 11.820$	$p_2 = .0533$
		$x_3 = 72.397$	$p_3 = .0007$

The arithmetic solution given can easily be improved: taking p_3 and x_3 fixed, one might replace x_1 and x_2 by the optimal diatomic feasible distribution for the remaining problem. By this method an optimal value of $\pi_S^* = 0.108535$ was found.

Which of the possibly many local maxima is found depends strongly both on the chosen starting point P^0 and on the value of h. The latter plays an unexpectedly and annoyingly important role, as can be seen from the following table for values of h and optima attained for the same example given above:

Span h	Optimum reached	Span h	Optimum reached
1.0000	0.107165	0.5000	0.107697
0.9900	0.108198	0.4950	0.108319
0.9850	0.108493	0.4925	0.108495
0.9800	0.108191	0.4900	0.108253

Note that by replacing $h = 1$ by $h = .99$, a much better maximum stop-loss premium was attained than by just halving it. As with every gradient method, one may have good or bad luck: at $h = 0.4925$ it took 10 runs with different starting solutions to improve upon the solution found in the first run with $h = 0.985$.

In theory, the problem at hand is solved by the above method. It does, however, take a lot of time. Moreover, it has to be executed many times, with different starting values and span widths, in order to make sure that the maximum found is indeed close enough to the global optimum. In practice we think it is advisable to use the somewhat cruder but much faster approach proposed in KAAS (1991), refinied if needed by admitting fixed mass on 0 and/or d. The reason is that in practice the data on μ, σ^2 and λ are based only

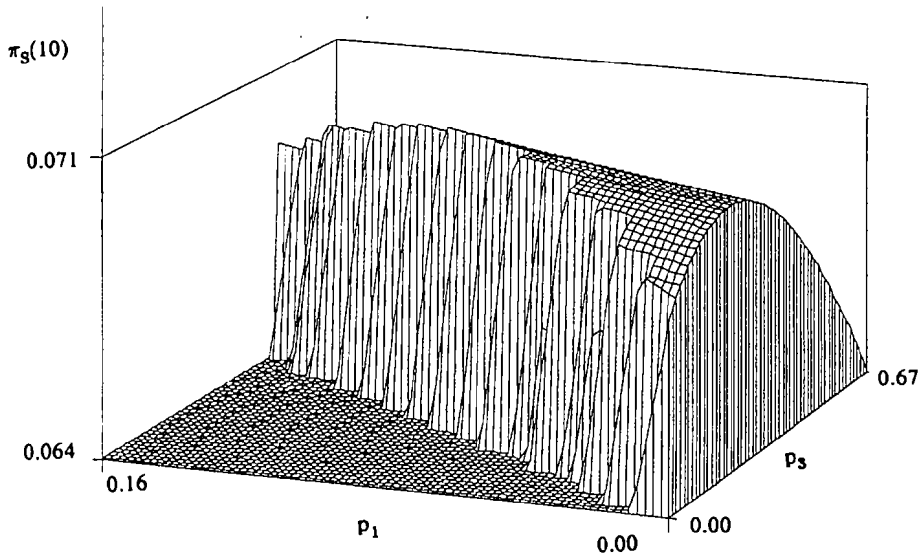

FIGURE 2. Stop-loss premium, expressed in p_1 and p_3, for the case $h = 1$, $n = 4$, $b = 4$, $m = 2$, $\mu = 3$, $\sigma^2 = 1$, $\lambda = 1$ and $d = 10$. The support is $\{0, 1, 2, 3, 4\}$.
Constraint $p_4 \geq 0$ is always fulfilled, $p_2 \geq 0$ gives $p_1 + p_3 \leq \frac{2}{3}$, $p_0 \geq 0$ gives $p_1 \leq \frac{1}{3}p_3$. Criterion values in the vertices: $p_1 = 0$, $p_3 = \frac{2}{3} \rightarrow .064$, $p_1 = \frac{1}{6}$, $p_3 = \frac{1}{2} \rightarrow .069$, and $p_1 = 0$, $p_3 = 0 \rightarrow .068$. There is a local maximum $p_1 = .0792$, $p_3 = .2377 \rightarrow .07044$. The global optimum is $p_1 = .0666$, $p_3 = .2515 \rightarrow .07064$.

on (often rather primitive) estimates. Furthermore, the compound Poisson model presupposes identically distributed claims, and independence of the claims and of the waiting times between them, which themselves must be exponential. All this might well not quite be fulfilled in practical situations. The method discussed in this paper has strengthened our belief that the diatomic maxima, though not always exactly optimal, are good enough.

REFERENCES

GOOVAERTS, M.J., KAAS, R., VAN HEERWAARDEN, A.E. and BAUWELINCKX, T. (1990) *Effective Actuarial Methods.* North-Holland, Amsterdam.

KAAS, R. (1991) The Schmitter problem and a related problem: A partial solution. *ASTIN Bulletin* **21**, 1, 133–146.

PRESS, W.H., FLANNERY, B.P., TEUKOLSKY, S.A. and VETTERLING, W.A. (1986) *Numerical Recipes, the art of scientific computing.* Cambridge University Press.

TAHA, H.A. (1987) *Operations Reseach: an introduction.* Mac Millan, New York, 4th edition.

ROB KAAS, MARLEEN VANNESTE AND MARC GOOVAERTS
Institute for Actuarial Science and Econometrics, Roetersstraat 11, NL-1018 WB Amsterdam.