ACTUARIAL RESEARCH CLEARING HOUSE
1994 VOL. 2

A Method For Calculating The Probabilities of Ruin By A Finite Number Of Claims Where The Process is Sparre-Andersen

by

William A. Bailey

Kemper National Insurance Companies
1 Kemper Drive
Long Grove, Illinois 60049-0001

Abstract

The Bivariate Numerical Generalized Convolution (Binugeco) Algorithm, presented in the Appendix of the ARCH (1993.1) paper "Six Bridges To Ψ's", can be applied to find the probability Ψ(u;n) of ruin by n claims for any Sparre-Andersen Process. We introduce a conjecture as to a method for estimating how much has to be added to Ψ(u;n) to obtain Ψ(u).

Table of Contents

Section 1. Introduction

This paper grew out of the work done in Bailey (1993b). That paper described six "bridges" to the probabilities of ruin where the process is Sparre-Andersen (see Thorin and Wikstad (1973). Although some univariate numerical convolutions were required for each of the bridges, three of the bridges involved Monte Carlo, two used further univariate convolutions and one used bivariate numerical generalized convolutions.

That paper indicated that some of the bridges might not be practical if the premium loading were very small or if the distributions were sufficiently pathological. In such cases the inutility of the bridges would become evident early in the calculations and other methods would have to be applied. The Committee on Papers suggested that, if any of the bridges were to be considered practical, it would have to be efficient enough to handle a Sparre-Andersen process which did not involve just a simple exponential distribution of waiting times between claims.

Part I of the current paper shows how Convolution Bridge #1 from Bailey (1993b) can be used to estimate for any Sparre-Andersen process the probability of ruin by a finite number of claims. Knowing the particular distribution of waiting times between claims and the particular distribution of size of individual claims, Part II of the current paper conjectures on how we might estimate the amount to add to

$\Psi(u;n)$ to obtain $\Psi(u)$, where $\Psi(u;n)$ is the probability of ruin by the n^{th} claim, and $\Psi(u)$ is the probability of ultimate ruin given an initial surplus of u. . A numerical example from Thorin and Wikstad (1973) is presented in Part III, along with some contrasts with methods of some other authors.

Section 2. Some Definitions

This paper focuses on the <u>Sparre-Andersen process</u>; that is, the model in which the waiting times $W_1, W_2, \ldots$ are independent and identically distributed as W, the size of claims $X_1, X_2, \ldots$ are independent and identically distributed as X, and $X_1, X_2, \ldots, W_1, W_2, \ldots$ are mutually independent.

Let c be the constant rate (per unit of time) at which premiums are received. Let W_1 be the waiting time until claim number 1; and, let W_j (j=2,3,...)be the waiting time between claim number j-1 and claim number j. Let X_k (k=1,2,...) be the size of claim number k. Let V_i be the amount in the fund just after claim number i, having started with an initial surplus of u. Then

$$V_i = u + c \cdot \sum_{j=1}^{i} W_j - \sum_{k=1}^{i} X_k .$$

We are concerned with the event that ruin occurs; that is, that $V_i < 0$ for some i=1,2,....

It is well known that a compound Poisson process involves an exponential distribution of waiting times. We will be requiring that the waiting times between claims be independent and identically distributed, but not necessarily exponentially.

We are requiring that E[X] be nonnegative, but X need not take on only nonnegative values,

Let the maximal aggregate loss random variable L_n be defined as

$$L_n = \max(0, -V_1, -V_2, \ldots, -V_n).$$

Given an initial surplus of u let $\Psi(u;n)$ be the probability of ruin by the n^{th} claim, and let $\Psi(u)$ be the probability of ultimate ruin.

Part I describes an almost exact method for determining $\Psi(u;n)$, using one algorithm (see the Appendix of Bailey (1993a)) for performing univariate numerical generalized convolutions and another algorithm (see Appendix #3 of Bailey (1993b)) for performing bivariate numerical generalized convolutions (binugecos).

Section 3. The Initial Discretization Procedure

In practice the distribution f_W of W and the distribution f_X of X would likely be obtained from empirical distributions implied by samples taken from relevant data for waiting time between claims and size of individual claim, respectively.

If f_W and f_X are to be continuous distributions, then each of these two distributions would be discretized. Although this initial discretization can be accomplished using discretization techniques such as those described by Dufresne and Gerber (1989) or Panjer (1986) , I have found the following approach works very well. These other discretization methods are vital to the calculation of $\Psi(u)$ in their respective papers; but they do not offer any advantage here, because of the way $\Psi(u)$ will be determined.

We consider f_X, although the procedure would be analogous for f_W.

Partition the x-axis into n intervals $[h_i, h_{i+1}]$ $(i=0,1,2,\ldots,n)$ where h_0 and h_n are chosen so that (where F_X is the cumulative distribution function),

$$F_X(h_0) \text{ is } < \varepsilon$$

and

$$F_X(h_n) \text{ is } > 1 - \varepsilon,$$

where n is chosen to be a large positive integer, say 10^5, and ε is chosen to be a small positive real number, say 10^{-12}. Often, the intervals are chosen to be of equal length, but <u>log intervals</u> are often useful; that is, where the h_i's are defined by

$$\log_{10}(h_i) = \log_{10}(h_0) + i \cdot \frac{\log_{10}(h_n) - \log_{10}(h_0)}{n} \qquad i=0,1,2,\ldots,n).$$

Sometimes, square-root intervals are useful; that is, where the h_i's are defined by

$$\sqrt{h_i} = \sqrt{h_0} + i \cdot \left(\frac{\sqrt{h_n} - \sqrt{h_0}}{n}\right) \qquad (i=1,2,\ldots,n)$$

Log intervals tend to produce finer results for smaller values of the random variables, whereas square-root intervals tend to produce finer results for larger values of the random variables.

The preliminary discretized form would be

$$\mathbf{f}_X = \left[\frac{\int_{h_{i-1}}^{h_i} x \cdot \mathbf{f}_X(x)dx}{\int_{h_{i-1}}^{h_i} \mathbf{f}_X(x)dx}, \int_{h_{i-1}}^{h_i} \mathbf{f}_X(x)dx \right]_{i=1,2,\ldots,n}$$

which is then convoluted with the degenerate distribution

$$f_O = [\ 0,\ 1\ ]$$

to generate the distribution $f_{X+O} = f_X * f_O$. The symbol * between two distributions is being used to indicate the convolution (for sums) of the two distributions. The purpose of convoluting with [0,1] is twofold; namely, (1) to be able to represent the distribution in a number of lines which is small enough to permit further convolutions being performed in a reasonable time, and (2) to retain accurately the first three moments of the distribution.

The probabilities in f_{W+O} and f_{X+O} are then normalized to sum to unity. Since the real objective of this paper is to handle the type of situation described in the first paragraph of this Section, we will not be concerned with any error introduced by this normalization.

Section 4. Generating The Distribution f_H

Let $f_H = f_{(1+\vartheta)\cdot W-X} = f_{(1+\vartheta)\cdot W+O} * f_{-X+O}$ where ϑ is the premium loading, the distributions $f_{(1+\vartheta)\cdot W+O}$ and f_{-X+O} are discretized as described in Section 3, and the symbol * between two distributions indicates the distribution generated by convoluting the two distributions.

The indicated convolution is, once again, performed by the univariate numerical convolution algorithm described in the Appendix of Bailey (1993a) or in Appendix #2 of Bailey (1993b). The parameters for that convolution can be chosen to be the same nax, type of mesh intervals and ε as was used in the convolution in Section 3. Although use of log intervals or square-root intervals (see Section 3) would at first glance appear unusable here because $H = (1+\vartheta)\cdot W-X$ takes on negative (as well as nonngegative) values, the convolution is performed in two pieces, namely, two partial distributions are generated: $\mathbf{f}^{-}_{-H}$ for $H <$ 0 and $\mathbf{f}^{+}_{H}$ for $H >= 0$, determining square-root intervals for each separately. The term partial distribution is being used to remind us that the sum of the probabilities in the distribution is far less than unity.

Section 5. Generating the Distributions $\mathbf{f}_{V_n,L_n}$

Then $V_1 = H$ and $L_1 = \max(0,-H)$. The distribution of V_1 is the concatenation of the partial distribution $\mathbf{f}^{-}_{-H}$ for $H<0$ and the partial distribution $\mathbf{f}^{+}_{H}$ for $H>=0$, each of the partial distributions having been generated by the method described in Section 4.

For n=1 we have $(V_1,L_1) = (H,\max(0,-H))$, so the bivariate distribution of (V_1,L_1) can be obtained by simple transforming the univariate distribution of H into the distribution of $(H, \max(0,-H))$.

For $n>1$, the bivariate distribution of (V_n, L_n) could be obtained recursively by performing bivariate numerical generalized convolutions of the bivariate distribution of (V_{n-1}, L_{n-1}) and the univariate distribution of H, using the formulas $V_n = V_{n-1} + H$ and $L_n = \max(-(V_{n-1}+H), L_{n-1})$.

However, this procedure may not be practical if n is large. So, we adopt the following alternative procedure.

For $n = 2^m$ with m any positive integer, the bivariate distribution of (V_n, L_n) can be obtained by recursively performing bivariate numerical generalized convolutions of the bivariate distribution of $(V_{n/2}, L_{n/2})$ and the bivariate distribution of $(\tilde{V}_{n/2}, \tilde{L}_{n/2})$, using the formulas $V_n = V_{n/2} + \tilde{V}_{n/2}$ and $L_n = \max(-(V_{n/2}-\tilde{L}_{n/2}) , L_{n/2})$ where $(\tilde{V}_{n/2}, \tilde{L}_{n/2})$ and $(V_{n/2}, L_{n/2})$ are independent and identically distributed. The formula for L_n follows from the fact that the maximal aggregate loss just after the n^{th} claim is the larger of (a) the maximal aggregate loss $L_{n/2}$ during the period of the first n/2 claims and the negative of (b) the fund $V_{n/2}$ just after the first n/2 claims, decreased by the maximal aggregate loss $\tilde{L}_{n/2}$ during the period of the next n/2 claims.

The indicated bivariate numerical generalized convolutions (referred to hereafter as binugecos) can be performed by the algorithm described in Appendix #3 of Bailey (1993b).

Section 6. Application of the Bivariate Numerical Generalized Convolution (Binugeco) Algorithm

For the binugecos indicated in Section 5, use of log or square-root intervals in the y-direction causes no special problems because L_n is never less than zero. (Zero is treated as a special degenerate mesh interval.) However, use of log or square-root intervals in the x-direction causes the same type of problem which was addressed in Section 4, because V_n takes on negative (as well as positive) values. Recalling the formulas from Section 5

$$V_n = V_{n/2} + \tilde{V}_{n/2}$$

and

$$L_n = \max(-(V_{n/2}-\tilde{L}_{n/2}) , L_{n/2})$$

we performed each binugeco in six pieces, namely, we generated the six partial distributions:

$$f^{(1)}_{-V_n,L_n} = f^{-}_{-V_{n/2},L_{n/2}} \bullet f^{-}_{-\tilde{V}_{n/2},L_{n/2}}$$

where $V_n < 0$, $V_{n/2} < 0$ and $\tilde{V}_{n/2} < 0$;

$$f^{(2)}_{-V_n,L_n} = f^{-}_{-V_{n/2},L_{n/2}} \bullet f^{+}_{\tilde{V}_{n/2},L_{n/2}}$$

where $V_n < 0$, $V_{n/2} < 0$ and $\tilde{V}_{n/2} > 0$;

$$f^{(3)}_{-V_n,L_n} = f^{+}_{V_{n/2},L_{n/2}} \bullet f^{-}_{-\tilde{V}_{n/2},L_{n/2}}$$

where $V_n < 0$, $V_{n/2} >= 0$ and $\tilde{V}_{n/2} < 0$;

$$f^{(4)}_{V_n,L_n} = f^{-}_{-V_{n/2},L_{n/2}} \bullet f^{+}_{\tilde{V}_{n/2},L_{n/2}}$$

where $V_n >= 0$, $V_{n/2} < 0$ and $\tilde{V}_{n/2} >= 0$;

$$f^{(5)}_{V_n,L_n} = f^{+}_{V_{n/2},L_{n/2}} \bullet f^{-}_{-\tilde{V}_{n/2},L_{n/2}}$$

where $V_n >= 0$, $V_{n/2} >= 0$ and $\tilde{V}_{n/2} < 0$;

$$f^{(6)}_{V_n,L_n} = f^{+}_{V_{n/2},L_{n/2}} \bullet f^{+}_{\tilde{V}_{n/2},L_{n/2}}$$

where $V_n >= 0$, $V_{n/2} >= 0$ and $\tilde{V}_{n/2} >= 0$;

determining suare-root intervals for each partial distribution separately. The superscripts in ()'s on **f** here are simply to identify the distributions in the following merges:

$f^{(1)}_{-V_n,L_n}$ and $f^{(2)}_{-V_n,L_n}$ and $f^{(3)}_{-V_n,L_n}$ are concatenated or merged together to form

$$f^{-}_{-V_n,L_n} \text{ where } V_n < 0;$$

$f^{(4)}_{V_n,L_n}$ and $f^{(5)}_{V_n,L_n}$ and $f^{(6)}_{V_n,L_n}$ are concatenated or merged together to form

$$f^{+}_{V_n,L_n} \text{ where } V_n >= 0.$$

The "merging" of the bivariate distributions can be done using the mesh rectangles described in Section 6 and in Appendix #3 of Bailey (1993b). This saves some computer time when we proceed to perform the next binugecos.

Part II. A Conjecture As To A Method For Determining The Probability of Ultimate Ruin

Section 7. Obtaining $\Psi(u)$

The Steps in this Section relate to the following two inequalities. These inequalities are presented prematurely as an aid to understanding why we perform the Steps described later in this Section.

$\Psi(u)$ Period:

$$<= 1 - \Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_0)} <= u\} + \Delta_0$$ 1st $\hat{n}$ claims

$$+ (\Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_0)} <= u\} - \Delta_0)$$
$$\cdot (1 - \Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_1)} <= \hat{l}_{2^{\hat{n}}}^{(\varepsilon_0)} \} + \Delta_1)$$ 2nd $\hat{n}$ claims

$$+ (\Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_0)} <= u\} - \Delta_0)$$
$$\cdot (\Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_1)} <= \hat{l}_{2^{\hat{n}}}^{(\varepsilon_0)} \} - \Delta_1)$$
$$\cdot (1 - \Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_2)} <= \hat{l}_{2^{\hat{n}}}^{(\varepsilon_1)} \} + \Delta_2)$$ 3rd $\hat{n}$ claims

$+ \ldots$

$$<= \Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_0)} > u \} + \Delta_0$$

$$+ \Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_1)} > \hat{l}_{2^{\hat{n}}}^{(\varepsilon_0)} \} + \Delta_1$$

$$+ \Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_2)} > \hat{l}_{2^{\hat{n}}}^{(\varepsilon_1)} \} + \Delta_2$$

$+ \ldots$

Breaking this latter series into two series we have:

(Series #1):

$$\Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_0)} > u \}$$
$$+ \sum_{i=1}^{\infty} \Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_i)} > \hat{l}_{2^{\hat{n}}}^{(\varepsilon_{i-1})} \}$$

and

(Series #2):

$$\Delta_0 + \sum_{i=1}^{\infty} \Delta_i$$

For Series #1 we look at the first 13 terms being summed and make a minimal assumption for the sum of the terms beyond that. See Section 8 below for numerical results.

For Series #2 there can be no more than (nax·nay) lines in one distribution (actually, (nax+1)·(nay+1) because 0 is handled as a degenerate interval for each of the two univariate variables). One binugeco could result in discarding no more than $((nax+1)\cdot(nay+1))^2$ probability products, each probability product not greater than ε_i. The number of convolutions being performed in one sequence of binugecos is $6\cdot\hat{n}$; 6 is the number of partial binugecos being performed for each of the desired binugecos. Thus we have a usable upper bound for Series #2; namely,

$$\sum_{i=1}^{\infty} \Delta_i \quad <= \quad 6\cdot\hat{n}\cdot((nax+1)\cdot(nay+1))^2\cdot\sum_{i=1}^{\infty} \varepsilon_i$$

where $\sum_{i=1}^{\infty} \varepsilon_i$ will, because of the way the ε_i's will be defined at the beginning of Step 7.3, be found to be small enough for the purposes of establishing an upper bound to $\Psi(u)$.

The Δ_m's, although they will be well-defined in the following Steps, may not be easy to compute in the absence of special software to handle arbitrary precision arithmetic. Fortunately, we will not have to calculate them.

Step 7.1 Determining $\hat{n}$

Let nax=32 and nay=32, say. Choose $\varepsilon_{-1}=10^{-299}$ (or the smallest positive real number which your computer will handle without underflowing), and perform the sequence of binugecos indicated in Sections 5 and 6 for $n=2,2^2,2^3,\ldots$. For each value of n, let $\hat{v}_{2^n}^{(\varepsilon_{-1})}$ be the smallest probable value of $V_{2^n}^{(\varepsilon_{-1})}$ in the marginal distribution $f_{V_{2^n}}^{(\varepsilon_{-1})}$ and let $\hat{l}_{2^n}^{(\varepsilon_{-1})}$ be the largest probable value of $L_{2^n}^{(\varepsilon_{-1})}$ in marginal distribution $f_{L_{2^n}}^{(\varepsilon_{-1})}$. Let $\hat{n}$ be the smallest value of n for which $\hat{v}_{2^n}^{(\varepsilon_{-1})}$ is greater than $\hat{l}_{2^n}^{(\varepsilon_{-1})}$; $\hat{n}$ is determined as the binugecos are performed and the minimum and maximum values of $V_{2^n}^{(\varepsilon_{-1})}$ and $L_{2^n}^{(\varepsilon_{-1})}$, respectively, in $f_{V_{2^n},L_{2^n}}^{(\varepsilon_{-1})}$ are observed. These marginal distributions are taken from the bivariate distribution $f_{V_{2^n},L_{2^n}}^{(\varepsilon_{-1})}$. Note that, although the probable values $\hat{v}_{2^n}^{(\varepsilon_{-1})}$ can be positive or negative, the negative values will disappear as n becomes larger; and, in fact, the probable values of $\hat{v}^{(\varepsilon_{-1})}_{2^n}$ will be greater than the probable values of $\hat{l}^{(\varepsilon_{-1})}_{2^n}$ for sufficiently large values of n. See Appendix #1 - Assertions in Bailey (1993b) for proof. The superscripts in parentheses are simply to indicate which ε was used in doing the sequence of binugecos. Thus, from this Step we have determined $\hat{n}$.

If we continued to generate $f^{(\varepsilon_m)}_{V_{2^n},L_{2^n}}$ for n = $\hat{n}$+1, $\hat{n}$+2, ..., we would soon find that the marginal distribution $f^{(\varepsilon_m)}_{L_{2^n}}$ becomes invariant. And, it would be tempting to simply use that invariant distribution to calculate $\Psi(u) = 1 - \Pr\{l<u\}$. However, we must consider the total of the probability products which we would be discarding. So, we proceed to Steps 7.2 and 7.3.

<u>Step 7.2</u> <u>Generating $f^{(\varepsilon_0)}_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}$</u>

Let nax=64 and nay=64, say, to produce finer distributions during the period of first $2^{\hat{n}}$ claims, where $\hat{n}$ was determined in Step 7.1. We are using larger values of nax and nay (than in the previous Step), in order to increase the accuracy of the results. Choose $\varepsilon_0=10^{-20}$ and perform the sequence of binugecos indicated in Sections 5 and 6 for $n=2,2^2,2^3,\dots,2^{\hat{n}}$. Let $\hat{v}^{(\varepsilon_0)}_{2^{\hat{n}}}$ be the smallest probable value of $V^{(\varepsilon_0)}_{2^{\hat{n}}}$ in the marginal distribution $f_{V_{2^{\hat{n}}}}^{(\varepsilon_0)}$ associated with the distribution $f_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}^{(\varepsilon_0)}$. Let $\hat{l}_{2^{\hat{n}}}^{(\varepsilon_0)}$ be the largest *probable* value of $L_{2^{\hat{n}}}^{(\varepsilon_0)}$ in the marginal distribution $f_{L_{2^{\hat{n}}}}^{(\varepsilon_0)}$ associated with the distribution $f_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}^{(\varepsilon_0)}$.

Let $\Delta_0 = 1 - \sum_{\text{probable } l\text{'s}} f_{L_{2^{\hat{n}}}}^{(\varepsilon_0)}(l)$.

Δ_0 is the sum of the probability products discarded in the $6\cdot\hat{n}$ binugecos performed in this Step.

<u>Step 7.3</u> <u>Generating</u> $f_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}^{(\varepsilon_m)}$

Our plan is to use $\varepsilon_m = 10^{-20\cdot m}$ during the m^{th} period of $2^{\hat{n}}$ claims, where m=1,2,.... The sequence $\{\varepsilon_m\}$ is designed so that $\sum_{m=1}^{\infty}\varepsilon_m$ will be small.

Second Period of $2^{\hat{n}}$ Claims:

Let nax=32 and nay=32. Choose $\varepsilon_1=10^{-40}$ and perform the sequence of binugecos indicated in Sections 5 and 6 for $n=2,2^2,2^3,\ldots 2^{\hat{n}}$. At this point we will have generated the distribution $f_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}^{(\varepsilon_1)}$, and we now describe how this distribution will be used.

Assume that the claim number $2^{\hat{n}}$ has just occurred. Focus on the period of the next $2^{\hat{n}}$ claims. Let $\hat{v}^{(\varepsilon_1)}_{2^{\hat{n}}}$ be the smallest probable value of $V^{(\varepsilon_1)}_{2^{\hat{n}}}$ in the marginal distribution $f^{(\varepsilon_1)}_{V_{2^{\hat{n}}}}$ associated with the distribution $f^{(\varepsilon_1)}_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}$; and, let $\hat{l}^{(\varepsilon_1)}_{2^{\hat{n}}}$ be the largest probable value of $L^{(\varepsilon_1)}_{2^{\hat{n}}}$ in the marginal distribution $f^{\varepsilon_1)}_{L_{2^{\hat{n}}}}$ associated with the distribution $f^{(\varepsilon_1)}_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}$.

If $L^{(\varepsilon_1)}_{2^{\hat{n}}} < \hat{l}^{(\varepsilon_0)}_{2^{\hat{n}}}$ and $\hat{l}^{(\varepsilon_0)}_{2^{\hat{n}}} < \hat{v}^{(\varepsilon_0)}_{2^{\hat{n}}}$, then

(1) $L^{(\varepsilon_1)}_{2^{\hat{n}}} < \hat{v}^{(\varepsilon_0)}_{2^{\hat{n}}}$ and

(2) ruin cannot occur during the 2^{nd} period of $2^{\hat{n}}$ claims.

This implies that

$Pr\{L^{(\varepsilon_1)}_{2^{\hat{n}}} < \hat{l}^{(\varepsilon_0)}_{2^{\hat{n}}}\}$ is <= the probability of <u>not</u> going ruin during the

period of the second $2^{\hat{n}}$ claims,

and

$Pr\{L^{(\varepsilon_1)}_{2^{\hat{n}}} >= \hat{l}^{(\varepsilon_0)}_{2^{\hat{n}}}\}$ is

>= the probability of going ruin during the

period of the second $2^{\hat{n}}$ claims. (Statement I)

We will see that $\hat{l}^{(\varepsilon_0)}_{2^{\hat{n}}} < \hat{v}^{(\varepsilon_0)}_{2^{\hat{n}}}$ in the table of Intermediate Numerical Results in Section 8.

Let

$$\Delta_1 = 1 - \sum_{\text{probable } l\text{'s}} f^{(\varepsilon_1)}_{L_{2^{\hat{n}}}}(l)$$

where $f^{(\varepsilon_1)}_{L_{2^{\hat{n}}}}$ is the marginal distribution of $L^{(\varepsilon_1)}_{2^{\hat{n}}}$ associated with the distribution, say $f^{(\varepsilon_1)}_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}$, generated using ε_1. Δ_1 is the sum of the probability products discarded in performing the binugecos.

m^{th} Period of $2^{\hat{n}}$ Claims:

Let nax=32 and nay=32. For each m (m=2,...,13), choose $\varepsilon_m = 10^{-m\cdot 20}$ and perform the sequence of binugecos indicated in Sections 5 and 6 for $n=2,2^2,2^3,\ldots 2^{\hat{n}}$. At this point we will have generated the distributions $f^{(\varepsilon_m)}_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}$ for each value of m=2,3,...,13. We now describe how these distributions will be used.

Assume that the claim number $(m-1)\cdot 2^{\hat{n}}$ has just occurred. Focus on the m^{th} period of $2^{\hat{n}}$ claims. For each m=2,3,....,13, let $\hat{v}^{(\varepsilon_m)}_{2^{\hat{n}}}$ be the smallest probable value of $V^{(\varepsilon_m)}_{2^{\hat{n}}}$ in the marginal distribution $f^{(\varepsilon_m)}_{V_{2^{\hat{n}}}}$

associated with the distribution $f^{(\varepsilon_m)}_{V_{2^{\hat{n}}}, L_{2^{\hat{n}}}}$; and, let $\hat{l}^{(\varepsilon_m)}_{2^{\hat{n}}}$ be the largest probable value of $L^{(\varepsilon_m)}_{2^{\hat{n}}}$ in the marginal distribution $f^{(\varepsilon_m)}_{L_{2^{\hat{n}}}}$ associated with the distribution $f^{(\varepsilon_m)}_{V_{2^{\hat{n}}}, L_{2^{\hat{n}}}}$.

If $L^{(\varepsilon_m)}_{2^{\hat{n}}} < \hat{l}^{(\varepsilon_{m-1})}_{2^{\hat{n}}}$ and $\hat{l}^{(\varepsilon_{m-1})}_{2^{\hat{n}}} < \hat{v}^{(\varepsilon_{m-1})}_{2^{\hat{n}}}$,

then

(1) $L^{(\varepsilon_m)}_{2^{\hat{n}}} < \hat{v}^{(\varepsilon_{m-1})}_{2^{\hat{n}}}$ and

(2) ruin cannot occur during the m^{th} period of $2^{\hat{n}}$ claims.

This implies that

$\Pr\{L^{(\varepsilon_m)}_{2^{\hat{n}}} < \hat{l}^{(\varepsilon_{m-1})}_{2^{\hat{n}}}\}$ is <= the probability of <u>not</u> going ruin during the period of the second $2^{\hat{n}}$ claims,

and

$\Pr\{L^{(\varepsilon_m)}_{2^{\hat{n}}} >= \hat{l}^{(\varepsilon_{m-1})}_{2^{\hat{n}}}\}$ is

>= the probability of going ruin during the period of the second $2^{\hat{n}}$ claims. (Statements II)

We will see that $\hat{l}^{(\varepsilon_m)}_{2^{\hat{n}}} < \hat{v}^{(\varepsilon_m)}_{2^{\hat{n}}}$ in the table of Intermediate Numerical Results in Section 8.

Let

$$\Delta_m = 1 - \sum_{\text{probable } l\text{'s}} f^{(\varepsilon_m)}_{L_{2^{\hat{n}}}}(l)$$

where $f_{L_{2^{\hat{n}}}}^{(\varepsilon_m)}$ is the marginal distribution of $L_{2^{\hat{n}}}^{(\varepsilon_m)}$ associated with the distribution, say $f_{V_{2^{\hat{n}}},L_{2^{\hat{n}}}}^{(\varepsilon_m)}$, generated using ε_m. For each value of m, Δ_m is the sum of the probability products discarded in performing the binugecos.

Statement I and Statements II above support the use of the sum of Series #1 at the beginning of this Section (in addition to the sum of Series #2) as an upper bound for $\Psi(u)$. Section 8 shows a numerical example, including $\Psi(u;2^{26})$ for four illustrative values of u; and, an estimate for $\Psi(u)-\Psi(u;2^{26})$. The Table labelled Intermediate Numerical Results in Section 8 may be helpful toward visualizing how and why the conjectured method of the current Section works.

Section 8. A Numerical Example

We have chosen to use the methods described above to calculate $\Psi(u)$ for u=0, 100, 1000 and 10000 for c=1.10 in Table 8 of Thorin and Wikstad (1973):

[Cumulative] Distribution of Interclaim times is

$$K(t) = 1 - 0.25 \cdot e^{-0.4 \cdot t} - 0.75 \cdot e^{-2 \cdot t}$$

[Cumulative] Distribution of Individual Claim Sizes is

$$P[y] = \sum_{v=1}^{5} a_v \cdot (1 - e^{-\alpha_v \cdot y})$$

where

v	a_v	α_v
1	0.6635948	3.675472
2	0.3114878	0.7116063
3	0.02405664	0.09447445
4	0.0008425574	0.009322980
5	0.00001823254	0.0004965620

Thorin & Wikstad claim that "In general the number of correct decimals are four but in most cases even five." Also, "The [cumulative] distribution functions have mean values equal to one."

In performing the initial discretization described in Section 3 square-root intervals were used, together with $\varepsilon=10^{-12}$. Before convoluting with [0,1] the distribution of waiting times had 82953 lines and the distribution of individual claim sizes had 82882 lines.

In performing the univariate numerical convolutions for sums in Section 4 and the bivariate numerical generalized convolutions in Section 5 and Section 7, square-root intervals were used. We used $\varepsilon=10^{-299}$ and 10^{-20} in Step 7.1 and 7.2, respectively; and $\varepsilon=10^{-40}, \ldots, 10^{-280}$ in Step 7.3. Univariate numerical convolutions were performed using nax=1000 and $\varepsilon=10^{-20}$.

Steps 7.1, 7.2 and 7.3 yielded the following:

Intermediate Numerical Results

i	ε_i	$\hat{v}_{2^{\hat{n}}}^{(\varepsilon_i)}$	$\hat{l}_{2^{\hat{n}}}^{(\varepsilon_i)}$	$\Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_i)} > \hat{l}_{2^{\hat{n}}}^{(\varepsilon_{i-1})} \}$
0	10^{-20}	5,788,512	56,902	
1	10^{-40}	5,163,687	200,793	$5.1 \cdot 10^{-11}$
2	10^{-60}	4,887,162	290,832	$3.6 \cdot 10^{-37}$
3	10^{-80}	4,553,684	413,256	$1.1 \cdot 10^{-52}$
4	10^{-100}	4,200,316	439,186	$2.1 \cdot 10^{-76}$
5	10^{-120}	3,941,165	625,146	$3.5 \cdot 10^{-78}$
6	10^{-140}	3,766,885	662,402	$9.7 \cdot 10^{-114}$
7	10^{-160}	3,345,253	708,318	$5.1 \cdot 10^{-121}$
8	10^{-180}	3,118,569	831,957	$1.2 \cdot 10^{-136}$
9	10^{-200}	2,949,523	1,018,835	$1.2 \cdot 10^{-147}$
10	10^{-220}	2,635,582	926,384	0.0
11	10^{-240}	2,567,677	1,189,038	$3.6 \cdot 10^{-206}$ •
12	10^{-260}	2,276,673	1,179,588	0.0
13	10^{-280}	2,034,681	1 460,415	$1.4 \cdot 10^{-253}$ •
14	10^{-299}		1,548,109	

•$\Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_i)} > \hat{l}_{2^{\hat{n}}}^{(\varepsilon_{i-2})} \}$

In view of the above intermediate numerical results, we made the modest assumption for this particular example that

$$\frac{\Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_{i+1})} > l_{2^{\hat{n}}}^{(\varepsilon_i)} \}}{\Pr\{ L_{2^{\hat{n}}}^{(\varepsilon_i)} > l_{2^{\hat{n}}}^{(\varepsilon_{i-1})} \}} = 10^{-1} \text{ for } i=14,15,\ldots$$

and we obtain an estimated upper bound for the sum of Series #1; namely,

$5.2\cdot 10^{-11}+1.4\cdot 10^{-253}\cdot(\frac{1}{1-\frac{1}{10}})$, which is less than 10^{-10}.

($5.2\cdot 10^{-11}$ is clearly less than the sum of the first 12 entries in the rightmost column of the above table Intermediate Numerical Results.)

For Series #2 we have

$$\Delta_0 + \sum_{i=1}^{\infty}\Delta_i <= \Delta_0 + 6\cdot\hat{n}\cdot((nax+1)\cdot(nay+1))^2\cdot\sum_{i=0}^{\infty}\varepsilon_i$$

$= 10^{-7} + 6\cdot 26\cdot(33\cdot 33)^2\cdot(10^{-20}\cdot(1+10^{-20}+10^{-40}+\ldots))$

$< 10^{-6}$, by summing the geometric series, which we use as an upper bound for Series #2.

Thus, as we consider the 2nd, 3rd, ... periods of $\hat{n}$ claims, we find that the periods of $\hat{n}$ claims after the first $\hat{n}$ claims require that we add less than $2\cdot 10^{-6}$ to our estimate of $\Psi(u;\hat{n})$ to get an estimate of $\Psi(u)$.

Before the complications of Section 7 we have:

$\Psi(u;2^{26})$

nax·nay	u= 0	10^2	10^3	10^4	runtime*
32·32	.9469	.5174	.2072	.0059	4
64·64	.9374	.4858	.2105	.0072	72
96·96	.9353	.4789	.2036	.0081	216
Thorin#	.9341	.4803	.2041	.0081	1/60

\# Thorin's figures are for $\Psi(u)$.

* in hours

Relative Error In $\Psi(u;2^{26})$

nax·nay	u= 0	10^2	10^3	10^4	
32·32	.0137	.0772	.0152	.2716	
64·64	.0035	.0115	.0314	.1111	
96·96	.0013	.0029	.0024	.0000	

My computer running was done on a Gateway 486/25 or 486/33; theirs was done on a Control Data 6600.

Section 9. Contrasting Methods

(1) The method used in Thorin and Wikstad (1973)

(a) restricts attention mostly to the following classes of distribution functions:

interclaim time distributions of the form

$$1-\sum_{j=1}^{m} b_j \cdot e^{\beta_j \cdot t};$$

claim amount distributions of the form

$$1-\sum_{j=1}^{m} a_j \cdot e^{\alpha_j \cdot y};$$

(b) the solution is obtained using funtions of a complex variable;

(c) apparently requires some trial-and-error, since "the algorithm is operating well only if the starting value is fairly good" and "a starting value has to be carefully chosen".

(d) calculates probabilities of ultimate ruin.

My method for calculating the probability of ruin by a finite number of claims does not restrict the form of the (iid) interclaim time distributions or the (iid) claim amount distributions, uses only real variables, and requires no trial-and-error.

Their computer runs took only one minute per table on a CDC 6600, whereas my method requires several hours on a Gateway 486/25 or 486/33. (A Hewlitt Packard 720 might take only $\frac{1}{8}$ as long as the hours shown in the above table, assuming enough RAM memory to avoid paging.)

(2) Dufresne and Gerber (1989) solutions require that the process be compound Poisson. If the claim amounts are restricted to be positive, they (and Panjer (1986)) use the compound geometric distribution, since then

$\Psi(0)$ is known to be $\frac{1}{1+\vartheta}$ and

the distribution of loss at time of first claim

is known to have probability density function

$$f_L(y) = \frac{1-P(y)}{\mu}.$$

If the claim amounts are allowed to be positive or negative but the claim amount distribution is of the form

$$\sum_{i=1}^{n} A_i\beta_i e^{-\beta_i \cdot y} \text{ with } y>0$$

$$\text{or} \quad \sum_{i=1}^{n} A_i\beta_i e^{-\beta_i \cdot (y+\tau)} \text{ with } y>-\tau$$

where the β_i's are positive parameters and

$A_1+A_2+...+A_n=1$

then they either solve an integral equation or use a Monte Carlo method to obtain the probability of ultimate ruin.

If the process is Sparre-Andersen but not compound Poisson, then my method applies; whereas their non-Monte-Carlo methods do not.

Although they give no computer timings, I assume that their non-Monte-Carlo methods are more efficient than mine; that is, that for a given amount of computing I assume they can obtain more accurate results than I can.

(3) Shiu (1989) used operational calculus to develop formulas for calculating the probability of ultimate ruin where the process is Compound Poisson. Operational calculus lies at the extreme opposite end of the spectrum from the rather pedestrian approach used in the current paper. Seah (1990) implements one of Shiu's formulas by incorporating an algorithm to reduce round-off error due to convolutions; and, he shows that one of Shiu's formulas is, and one is not, practical for computing.

Section 10. Conclusion

Where the Sparre-Andersen process is restricted to Compound poisson, methods exist (see Section 9) to calculate or estimate the probability of ultimate ruin and do so more efficiently (i.e. more accurately for a given amount of computing) than the methods of the current paper. On the other hand, where the Sparre-Andersen process is restricted to be other than Compound poisson and other than distributions handled by Thorin and Wikstad (1973), the method of the current paper for calculating the probability of ruin by a finite number of claims still applies. The amount of computer time required by the method of the current paper depends on the degree of accuracy sought.

There is no Sparre-Anderen process for which the probability ruin by a finite number of claims can not be obtained by the method of the current paper; and the current paper includes a conjecture as to a method for obtaining the probability of ultimate ruin for such processes.

Acknowledgments

My sincere appreciation to Larry Hickey of PolySystems, Inc. in Chicago for having focused my attention on Von Mises Theorem and having been involved in the design and development of earlier versions of COCONUT™ which is a "C" language computer package which implements the numerical generalized convolution algorithms, and Josh Zirin (formerly of Kemper Group) for assistance in implementing the Von Mises algorithm.

References

Bailey, W. A. (1993a) A Method For Determining Confidence Intervals For Trend. *TRANSACTIONS OF THE SOCIETY OF ACTUARIES, Volume XLIV*, pp 11-53.

Bailey, W. A. (1993b) Six Bridges to Ψ's. *ARCH 1993.1*, pp 143-227.

Gerber, H. U. and Dufresne, F. (1989) Three Methods To Calculate The Probability of Ruin. *ASTIN BULLETIN, Volume 19, No.1*, pp 71-90.

Panjer, H. H. (1986) Direct Calculation of Ruin Probabilities. *JOURNAL OF RISK*

Seah, E. S. (1990) Computing the Probability of Eventual Ruin. *TRANSACTIONS OF THE SOCIETY OF ACTUARIES, Volume XLII*, pp 421-431.

Shiu, E. W. (1989) Ruin Probability by Operational Calculus.

Insurance: Mathematics and Economics 8, 243-249.

Springer, M. D. (1979) *The Algebra of Random Variables*, John Wiley & Sons, Inc., New York, 269-270.

Thorin, O. and Wikstad, N. (1973) Numerical Evaluation Of Ruin Probabilities For A Finite Period. *ASTIN Bulletin* 57, Part 2, September 1973, 137-153.