THE CALCULATION OF AGGREGATE LOSS DISTRIBUTIONS FROM CLAIM SEVERITY AND CLAIM COUNT DISTRIBUTIONS

PHILIP E. HECKMAN
GLENN G. MEYERS

Abstract

This paper discusses aggregate loss distributions from the perspective of collective risk theory. An accurate, efficient and practical algorithm is given for calculating cumulative probabilities and excess pure premiums. The input required is the claim severity and claim count distributions.

One of the main drawbacks of the collective risk model is the uncertainty of the parameters of the claim severity and claim count distributions. Modifications of the collective risk model are proposed to deal with these problems. These modifications are incorporated into the algorithm.

Examples are given illustrating the use of this algorithm. They include (1) calculating the pure premium for a policy with an aggregate limit; (2) calculating the pure premium of an aggregate stop-loss policy for group life insurance; and (3) calculating the insurance charge for a multi-line retrospective rating plan, including a line which is itself subject to an aggregate limit.

1. INTRODUCTION

This paper discusses aggregate loss distributions from the perspective of collective risk theory. Our objective is to provide an efficient algorithm for calculating the cumulative probabilities and excess pure premium ratios for aggregate loss distributions in terms of the claim severity and claim count distributions. Examples illustrating the use of this algorithm will be given.

Aggregate loss distributions are playing an increasingly important role in the pricing of insurance coverages. The insurance buying public is becoming more sophisticated and is recognizing that it is to their advantage to absorb as much of their losses as they possibly can and to purchase excess insurance to cover the catastrophic losses. With the degree of competition that exists in the insurance marketplace, it is extremely important to obtain accurate estimates of the losses that could arise from such an insurance contract.

Aggregate loss distributions have been widely discussed in the insurance literature. Members of the Casualty Actuarial Society are familiar with the papers of Dorweiler [1], Valerius [2], Simon [3] and Hewitt [4]. The aggregate loss distributions in these papers are based on observed aggregate loss data of individual insureds. A problem with this approach is that to get a sufficient volume of data, one must combine the experience of insureds for which one would expect different aggregate loss distributions.

The use of collective risk theory provides an alternative to the above approach. Aggregate loss distributions are calculated in terms of the underlying claim severity and claim count distributions. Empirical data on claim severity and claim count distributions are, in many cases, readily available. Many feel that this approach is superior to observing actual aggregate losses because it makes more efficient use of available data. Much relevant detail is buried when one observes only aggregate loss data.

However, the collective risk model does have some drawbacks. There are problems involved in fitting a distribution to the claim count. For a given insured we get one measurement of the claim count per year. During the years that we get the measurements, the exposure of the insured is most likely changing. In addition, observations are clouded by the fact that we must estimate the number of claims which have been incurred but not reported. Because of these problems it is difficult to fit a distribution to the claim count. Often, we must assume a distribution (usually Binomial, Poisson or Negative Binomial) with the parameters selected by judgment.

While empirical claim severity distributions are readily available, there are still some formidable problems that must be solved. There is no consensus as to how claim severity distributions should be adjusted for inflation. If we try to minimize this problem by choosing a relatively recent claim severity distribution, we will understate the variance of the ultimate claim severity distribution. To see this, consider the following equation.

$$\text{Var}\,(Z) = E_R(\text{Var}\,(Z|R)) + \text{Var}_R\,(E(Z|R))$$

$$Z = \text{Ultimate Loss}$$

$$R = \text{Case Reserve}$$

When case estimates are set at the expected value of the ultimate payment, the variance of the immature distribution will be $\text{Var}_R(E(Z|R))$. The variance of the ultimate claim severity distribution will be greater! Great care must be exercised in selecting the ultimate loss distribution. Methods of solving this problem can be found in the literature. [5][6].

Another problem with the collective risk model is that the calculation of the aggregate loss distribution has been very difficult. A great deal has been written about the various methods of solving this problem. We shall attempt to summarize these methods.

One general approach has been to calculate the moments of the aggregate loss distribution in terms of the moments of the claim severity distribution and the claim count distribution. One can then match the moments of the aggregate loss distribution with an assumed distribution. Probably the best known example of this approach is the Normal Power approximation [7]. However, the conditions required to insure the accuracy of this method can be very restrictive. Gary Venter uses the transformed Gamma distribution and obtains better results [8]. While it is easy to compute the results using these methods, one runs the risk of inaccuracies because the assumed distribution is not the same as that implied by the collective risk model.

A very popular method of calculating the aggregate loss distribution is by Monte Carlo simulation. Glenn Meyers has written an article illustrating this approach [9]. This method is easy to understand and can be very accurate. However, it currently requires a great deal of computer time.

A third method of calculating the aggregate loss distribution involves inverting its characteristic function. A recent article illustrating this approach was written by Dr. Shaw Mong [10]. This method requires that we have an explicit representation of the characteristic function of the claim severity distribution. Mong uses a shifted Gamma distribution to describe the claim severity distribution. Mong gives formulas for approximating other claim severity distributions with the shifted Gamma by matching the first three moments. The accuracy of this method depends upon how well the shifted Gamma distribution approximates the desired claim severity distribution.

A fourth method is the so-called "recursive method." This method assumes a discrete claim severity distribution. By choosing a large enough number of points for the claim severity distribution, one can obtain any desired degree of accuracy. For this reason, it has been called an "exact" method. This method requires far less computer time than Monte Carlo simulation. The recursive method is derived in papers by Ethan Stroh [11] and James Tilley [12] by inverting the Laplace transform of the aggregate loss distribution. Much of the mathematics involved is similar to that used in the characteristic function inversion method. Harry Panjer gives a derivation of the recursive method which does not involve inverting the Laplace transform [13].

The method described in this paper inverts the characteristic function of the aggregate loss distribution. Like the recursive method, it is an exact method. Its application goes beyond the recursive method in the following ways.

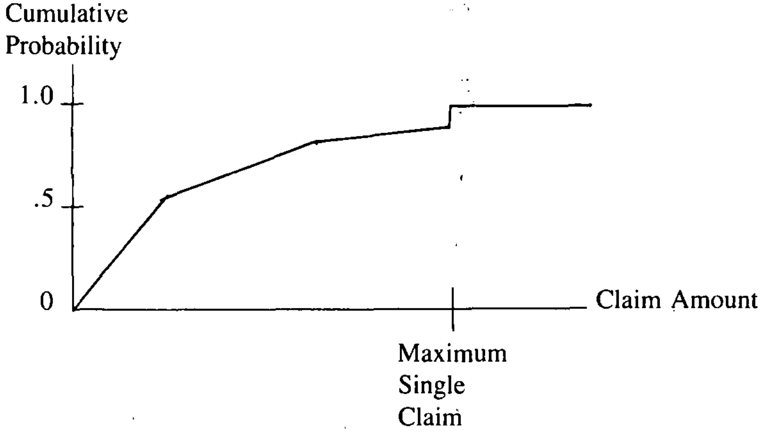
1. This method allows one to combine the aggregate loss distributions of several different lines into a composite aggregate loss distribution. This is necessary if one is to apply the results of the collective risk model to multi-line retrospective rating plans.
2. This method allows for parameter uncertainty in both the claim severity distribution and the claim count distribution. Glenn Meyers and Nathaniel Schenker have shown that allowing for parameter uncertainty significantly improves the fit of the collective risk model to empirical data [14]. It should be noted that Gary Venter's method of Reference [8] also allows for parameter uncertainty.
3. Philip Heckman and Phillip Norton have used the results of this paper to derive a method of selecting the specific and aggregate policy limits that minimize the variance of the retained losses while holding the cost of coverage constant [15].

In short, this method is applicable to a wide variety of insurance pricing problems. We include several examples which illustrate this.

The input required for this algorithm will be the claim count distribution and the claim severity distribution for each exposure class covered by the insurance contract. The claim count distribution can be either Binomial, Poisson or Negative Binomial. The cumulative claim severity distribution is assumed to be piecewise linear. We also allow the highest possible claim amount to occur with some non-zero probability. Figure 1 shows a cumulative distribution function that might typically be considered. Since most claim severity distributions applicable to the insurance business can be approximated to any desired degree of accuracy by a piecewise linear cumulative distribution, we feel we have a completely general method of performing these calculations.

It should be noted that these calculations will require a computer. With the nearly universal availability of computers, we do not consider this a drawback. We will warn the reader that the calculations are very complex, but, at the risk of being repetitious, we will stress the underlying concepts at every opportunity. This method is far more efficient than the more easily understood process of Monte Carlo simulation. Having fulfilled our duty to warn the reader, let us now proceed.

FIGURE 1



2. SUMMARY

We begin by giving a full description of the aggregate loss model. We will show how this distribution can be expressed empirically in terms of Monte Carlo simulation or analytically in terms of convolutions.

After reviewing some basic properties of complex numbers, we will introduce the characteristic function of a probability distribution. One of the remarkable properties of this complex-valued function is that the characteristic function of the convolution of two probability distributions is equal to the product of the characteristic functions of the two individual probability distributions. It is this property of characteristic functions that makes this method work. It is easier to multiply characteristic functions than it is to calculate convolutions by Monte Carlo simulation.

The next section will express the characteristic function of the aggregate loss distribution in terms of the claim count distribution and the characteristic function of the claim severity distribution. We will then derive formulas for the cumulative probabilities and the excess pure premiums for the aggregate loss distribution in terms of its characteristic function. These formulas involve improper integrals which can be evaluated using a Gaussian quadrature formula.

We then provide an analysis of the errors made in numerically evaluating the improper integrals. In some cases, the aggregate loss distribution is known. We test this algorithm by comparing the calculated results with known results. We also provide a comparison of the calculated results with results produced by Monte Carlo simulation.

Four examples illustrating the use of this algorithm will be given: (1) calculating the pure premium of a policy with aggregate limits; (2) calculating the pure premium of an aggregate stop-loss policy for group life insurance; (3) calculating the insurance charge for a retrospective rating plan involving two policies, one of which is subject to an aggregate limit; and (4) an example similar to (3) except that there is parameter uncertainty for the claim severity distribution.

3. THE COLLECTIVE RISK MODEL

Collective risk theory started by considering the generalized Poisson distribution. However, it soon became apparent that the assumptions of this distribution are violated for many applications. In this section we will discuss the assumptions of the generalized Poisson distribution and indicate some common violations of these assumptions. We will then state a version of the collective risk model that can deal with certain violations of these assumptions.

We start by considering the Poisson distribution. The assumptions underlying this distribution are as follows [16].

1. Claims occurring in two disjoint time intervals are independent.
2. The expected number of claims in a time interval (t_1, t_2) is dependent only on the length of the interval and not on the initial value t_1.
3. No more than one claim can occur at a given time.

There are situations when these assumptions are violated. We give three examples.

1. *Positive Contagion*
   A manufacturer can be held liable for defects in its products which, in many cases, are mass produced. A successful claim against the manufacturer may result in several other claims against the manufacturer. The notion that a higher than expected number of claims in an earlier period can increase the expected number of claims in a future period is what is called positive contagion.

2. *Negative Contagion*
   Consider a group life insurance policy. A death in an earlier period will reduce the expected number of deaths in a later period. One does not die twice. The notion that a higher than expected number of claims in an earlier period can decrease the expected number of claims in a future period is called negative contagion.

3. *Parameter Uncertainty*
   There are many cases when one feels that a Poisson distribution is appropriate, but one does not know the expected number of claims. Two options are available under these circumstances. The first option is to estimate the expected number of claims using historical experience. Parameter uncertainty can come from sample error. A second option is to use the average number of claims for a group of insureds that are similar to the insured under consideration. Parameter uncertainty arises if each member of the group expects to have a different number of claims.

   The effect of parameter uncertainty is similar to that of positive contagion. We give a heuristic argument for this which appeals to modern credibility theory. Suppose one observes n claims during a certain time period. Then one can estimate the number of claims, x, in a future period of equal length using the following formula.

   $$x = Z \cdot n + (1 - Z) \cdot E$$

   where E = Prior estimate
   Z = Credibility factor.

   Note that if the estimate of the expected number of claims is precise or the group of insureds is homogeneous, the credibility factor will be 0.

   If n is greater than expected, the number of claims expected in the future will be greater than the prior estimate for non-zero values of credibility.

   It should be emphasized that we are not arguing that claims in an earlier period will cause claims in a later period, as in the positive contagion case. We are stating only that the claim count distributions observed under the conditions of parameter uncertainty and positive contagion should be similar.

We now turn to specifying the claim count distributions we shall use for each of the above situations. We shall adopt the following notation.

n - A random variable denoting the number of claims
λ - The expected number of claims ($\lambda = E(n)$)
χ - A random variable with $E(\chi) = 1$ and Var $(\chi) = c$

Parameter uncertainty can be modeled by the following algorithm.

Algorithm 3.1

1. Select χ at random from the assumed distribution.
2. Select the number of claims, n, at random from a Poisson distribution with parameter $\chi\lambda$.

We have the following relationships.

$$E(n) = E(n|\chi) \cdot E(\chi) = \lambda \tag{3.1}$$

$$\begin{aligned} \text{Var}\,(n) &= E_\chi(\text{Var}\,(n|\chi)) + \text{Var}_\chi\,(E(n|\chi)) \\ &= E_\chi(\chi\lambda) + \text{Var}_\chi\,(\chi\lambda) \\ &= \lambda + c\lambda^2. \end{aligned} \tag{3.2}$$

If χ has a Gamma distribution, the claim count distribution described by Algorithm 3.1 is the Negative Binomial distribution [17]. We shall use the Negative Binomial distribution to model both the positive contagion and the parameter uncertainty cases.

We shall call the paramter c the contagion parameter for the claim count distribution. We should note that c has also been called the contamination parameter by some authors [18]. It should be noted that if $c = 0$, Algorithm 3.1 yields the Poisson distribution.

We shall use the Binomial distribution to model the negative contagion case. If m is the number of trials and p is the probability of success, we can formally define the contagion parameter to be equal to $-1/m$. Substituting this into Equation 3.2 yields the correct Binomial variance.

$$\text{Var}(n) = \lambda - \lambda^2/m = mp - m^2p^2/m = mp(1 - p)$$

While a negative contagion parameter makes no sense in terms of Algorithm 3.1, we shall see below that this is a very appropriate definition.

We now adopt the following notation.

z—A random variable denoting the amount of a claim
$S(z)$—The cumulative distribution function of z
x—A random variable denoting the aggregate loss for an insured

Aggregate losses can be generated by the following algorithm.

Algorithm 3.2

1. Select the number of claims, n, at random from the assumed claim count distribution.
2. Do the following n times.
   2.1 Select the claim amount, z, at random from the assumed claim severity distribution.
3. The aggregate loss amount, x, is the sum of all claim amounts, z, selected in step 2.1.

Let $F(x)$ denote the cumulative distribution function for the aggregate losses generated by Algorithm 3.2. We now give expressions for the mean and the variance of this distribution.

$$E(x) = E(n) \cdot E(z) = \lambda \cdot E(z) \tag{3.3}$$

$$\begin{aligned} \text{Var}\,(x) &= E_n(\text{Var}\,(x|n)) + \text{Var}_n\,(E(x|n)) \\ &= E_n(n \cdot \text{Var}\,(z)) + \text{Var}_n\,(n \cdot E(z)) \\ &= \lambda \cdot \text{Var}\,(z) + (\lambda + c\lambda^2) \cdot E^2\,(z) \\ &= \lambda \cdot E(z^2) + c\lambda^2 \cdot E^2\,(z) \end{aligned} \tag{3.4}$$

Implicit in the use of Algorithm 3.2 is the assumption that the claim severity distribution, $S(z)$, is known. In practice this distribution must be estimated from historical observations, or it must be simply assumed. Parameter uncertainty of the claim severity distribution can significantly affect the predictions of the collective risk model, and it should not be ignored. Our response to this problem is to make the simplifying (and we think reasonable) assumption that the shape of the distribution is known but there is uncertainty in the scale of the distribution.

More precisely, we specify parameter uncertainty of the claim severity distribution in the following manner. Let β be a random variable satisfying the conditions $E(1/\beta) = 1$ and $\text{Var}\,(1/\beta) = b$. We then model aggregate losses by the following algorithm.

Algorithm 3.3

1. Select the number of claims, n, at random from the assumed claim count distribution.
2. Select the scaling parameter, β, at random from the assumed distribution.

3. Do the following n times.
   3.1 Select the claim amount, z, at random from the assumed claim severity distribution.
4. The aggregate loss amount, x, is the sum of all claim amounts, z, divided by the scaling parameter, β.

Let $\mathcal{F}(x)$ denote the cumulative distribution function for the aggregate losses generated by Algorithm 3.3. Let $U(\beta)$ be the cumulative distribution function for the scaling parameter, β. Then the relationship between $\mathcal{F}(x)$ and $F(x)$ is given by the following equation.

$$\mathcal{F}(x) = \int_0^\infty F(\beta x)dU(\beta) \tag{3.5}$$

We now give formulas for the mean and the variance for the aggregate losses generated by Algorithm 3.3.

$$\begin{aligned} E(x) &= E_\beta(E(x|\beta)) \\ &= E_\beta(\lambda \cdot E(z)/\beta) \\ &= \lambda \cdot E(z) \cdot E(1/\beta) \\ &= \lambda \cdot E(z) \end{aligned} \tag{3.6}$$

$$\begin{aligned} \text{Var}\,(x) &= E_\beta(\text{Var}\,(x|\beta)) + \text{Var}_\beta\,(E(x|\beta)) \\ &= E_\beta[(\lambda \cdot E(z^2) + c\lambda^2 \cdot E^2(z))/\beta^2] + \text{Var}_\beta\,(\lambda \cdot E(z)/\beta) \\ &= (\lambda \cdot E(z^2) + c\lambda^2 \cdot E^2(z)) \cdot E(1/\beta^2) + \lambda^2 \cdot E^2(z) \cdot \text{Var}\,(1/\beta) \\ &= (\lambda \cdot E(z^2) + c\lambda^2 \cdot E^2(z)) \cdot (1+b) + \lambda^2 \cdot E^2(z) \cdot b \\ &= \lambda \cdot E(z^2)\,(1+b) + \lambda^2 \cdot E^2(z) \cdot (b+c+bc) \end{aligned} \tag{3.7}$$

In this paper, we shall assume that β has a Gamma distribution. We shall call b the mixing parameter. The mixing parameter is a measure of parameter uncertainty for the claim severity distribution.

It should be noted that we have chosen mathematically convenient distributions to model contagion and parameter uncertainty. We do not want to imply that these distributions are in any way the "correct" ones. Since parameter uncertainty is not directly observable, it is difficult to discover what the proper distribution should be. It should be noted that it is possible to infer the variance of the parameter uncertainty through the use of Equations 3.4 and 3.7 [14]. But until statistical methodology has advanced to the point where the proper distribution can be determined, it should be acceptable to use ones which are mathematically convenient.

4. CONVOLUTIONS

The above discussion provides a complete description of the aggregate loss model we use in this paper. Algorithms 3.2 and 3.3 provide the means to calculate the cumulative distribution by Monte Carlo simulation. Unfortunately this is a long and expensive process. We now begin to develop the mathematical tools necessary to derive a more efficient process.

Initially we will be concerned with the cumulative distribution function $F(x)$ which is described by Algorithm 3.2. We will then make use of Equation 3.5 to derive the cumulative distribution function $\mathcal{F}(x)$ described by Algorithm 3.3.

Let x be a random variable which has a distribution function $F(x)$. Similarly, let y be a random variable which has distribution function $G(y)$. Let $z = x + y$. Then the convolution of F and G, denoted by $F * G$ is the distribution function for z. We can express this analytically by the equation

$$(F * G)(z) = \int_0^z F(z-y)dG(y).$$

Let $S(z)$ be a claim severity distribution. Define

$$S^{0^*}(z) = \begin{cases} 0 \text{ if } z = 0 \\ 1 \text{ if } z > 0 \end{cases}$$

$$S^{n^*}(z) = (S^{(n-1)^*} * S)(z).$$

One can see that $S^{n^*}(z)$ is the distribution of the total amount of exactly n claims.

Algorithm 3.2 can be expressed in the following manner.

Algorithm 4.1

1. Select the claim count, n, at random.
2. Select the aggregate loss amount, x, from the distribution S^{n^*}.

We now give an analytical expression for this process. Let $F(x)$ denote the distribution function for the aggregate loss distribution. Let $P(n)$ denote the probability of exactly n claims. We then have

$$F(x) = \sum_{n=0}^{\infty} P(n) \cdot S^{n^*}(x). \tag{4.1}$$

5. CHARACTERISTIC FUNCTIONS

It may be helpful at this point to review some properties of complex numbers. A complex number, z, is one which can be written in the form

$$z = a + bi, \tag{5.1}$$

where a and b are real numbers and $i = \sqrt{-1}$. The number a is called the real part of z and b is called the imaginary part of z. Alternatively, z can be written in the form

$$z = re^{i\theta}, \tag{5.2}$$

where r is a nonnegative real number and θ is any real number; r is called the modulus of z, and θ is called the argument of z.

The equivalence of Equations 5.1 and 5.2 can be seen by using Euler's formula.

$$e^{i\theta} = \cos(\theta) + i \cdot \sin(\theta) \tag{5.3}$$

Using this formula it is not difficult to show that the following relationships hold.

$$r = \sqrt{a^2 + b^2} \tag{5.4}$$

$$\theta = \begin{cases} \arctan(b/a) \text{ if } a > 0 \\ \pi + \arctan(b/a) \text{ if } a < 0 \text{ and } b \geq 0 \\ \arctan(b/a) - \pi \text{ if } a < 0 \text{ and } b \leq 0 \\ \pi/2 \text{ if } a = 0 \text{ and } b > 0 \\ -\pi/2 \text{ if } a = 0 \text{ and } b < 0 \end{cases} \tag{5.5}$$

$$a = r\cos(\theta) \tag{5.6}$$

$$b = r\sin(\theta) \tag{5.7}$$

Having given a brief discussion of complex numbers, we define the characteristic function (or Fourier transform) of a cumulative distribution function F.

$$\phi_F(t) = E(e^{itx}) = \int_0^\infty e^{itx} dF(x) \tag{5.8}$$

Let F and G be two cumulative distribution functions.

$$\phi_{(F*G)}(t) = E(e^{itz}) = \int_0^\infty e^{itz} d(F * G) .$$

Since z is the sum of x and y, and x and y are independent, we have

$$\phi_{(F*G)}(t) = E_z(e^{itz}) = E_{x,y}(e^{it(x+y)}) = E_x(e^{itx})\,E_y(e^{ity}) = \phi_F(t)\phi_G(t).$$

Thus we have proved the following well known result.

$$\phi_{(F*G)}(t) = \phi_F(t)\,\phi_G(t) \tag{5.9}$$

As a consequence of this equation we have the following.

$$\phi_{S^{n*}}(t) = (\phi_S(t))^n \tag{5.10}$$

Combining Equations 4.1 and 5.10 we get the following expression for the characteristic function of an aggregate loss distribution, F.

$$\phi_F(t) = \sum_{n=0}^{\infty} P(n)(\phi_S(t))^n \tag{5.11}$$

As stated above, we assume that the claim severity distribution is piecewise linear. We now specify the mathematical form of the claim severity distribution, $S(z)$.

1. Let n be a nonnegative integer.
2. Let $0 \le a_1 < \ldots < a_n < a_{n+1}$.
3. Let p_k denote the probability that an individual loss is between a_k and a_{k+1}.
4. For $a_k < z < a_{k+1}$, the probability density of z is given by $d_k = p_k/(a_{k+1} - a_k)$.
5. The probability that a claim is equal to a_{n+1} is given by

   $$1 - \sum_{k=1}^{n} p_k.$$

   This allows us to describe the accumulation of claim values at the policy limit (a_{n+1}).

We now calculate the characteristic function of $S(z)$.

$$\phi_S(t) = \int_0^{\infty} e^{itz} dS(z)$$

$$\phi_S(t) = \sum_{k=1}^{n} \int_{a_k}^{a_{k-1}} d_k \cdot e^{itz} dz + \left(1 - \sum_{k=1}^{n} p_k\right) e^{ita_{n-1}}$$

Using Euler's formula (Equation 5.3) we continue.

$$\phi_S(t) = \frac{1}{t}\sum_{k=1}^{n} d_k(\sin(ta_{k+1}) - \sin(ta_k)) + \left(1 - \sum_{k=1}^{n} p_k\right)\cos(ta_{n+1})$$

$$+ i\frac{1}{t}\sum_{k=1}^{n} d_k(\cos(ta_k) - \cos(ta_{k+1})) + i\left(1 - \sum_{k=1}^{n} p_k\right)\sin(ta_{n+1})$$

Let $\tilde{h}(t)$ and $\tilde{k}(t)$ denote the real and imaginary parts of $\phi_S(t)$ respectively.

$$\tilde{h}(t) = \frac{1}{t}\sum_{k=1}^{n} d_k(\sin(ta_{k+1}) - \sin(ta_k)) + \left(1 - \sum_{k=1}^{n} p_k\right)\cos(ta_{n+1}) \quad (5.12)$$

$$\tilde{k}(t) = \frac{1}{t}\sum_{k=1}^{n} d_k(\cos(ta_k) - \cos(ta_{k+1})) + \left(1 - \sum_{k=1}^{n} p_k\right)\sin(ta_{n+1}) \quad (5.13)$$

We now turn to the problem of calculating the characteristic function of the aggregate loss distribution. Our main tool will be Equation 5.11.

Case 1 Binomial Distribution $\mathrm{P(n)} = \binom{\mathrm{m}}{\mathrm{n}} \mathrm{p^n} (1-\mathrm{p})^{\mathrm{m-n}}$

$$\phi_F(t) = \sum_{n=0}^{m} \binom{m}{n} p^n(1-p)^{m-n}(\phi_S(t))^n$$

$$\phi_F(t) = \sum_{n=0}^{m} \binom{m}{n} (p\phi_S(t))^n \cdot (1-p)^{m-n}$$

$$\phi_F(t) = (p\phi_S(t) + 1 - p)^m$$

$$\phi_F(t) = (1 + p(\phi_S(t) - 1))^m$$

If we make a change of notation and let $\lambda = mp$ and $c = -1/m$, we get

$$\phi_F(t) = (1 - c\lambda(\phi_S(t) - 1))^{-1/c}. \quad (5.14)$$

Case 2 Poisson Distribution $P(n) = \dfrac{e^{-\lambda}\lambda^n}{n!}$

$$\phi_F(t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} (\phi_S(t))^n$$

$$\phi_F(t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}(\lambda \cdot \phi_S(t))^n}{n!}$$

$$\phi_F(t) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda \cdot \phi_S(t))^n}{n!}$$

$$\phi_F(t) = e^{-\lambda} e^{\lambda \cdot \phi_S(t)}$$

$$\phi_F(t) = e^{\lambda \cdot (\phi_S(t) - 1)} \tag{5.15}$$

Case 3 Negative Binomial Distribution

$$P(n) = \binom{n + 1/c - 1}{n} (1 + c\lambda)^{-1/c} \left(\frac{c\lambda}{1 + c\lambda}\right)^n$$

$$\phi_F(t) = \sum_{n=0}^{\infty} \binom{n + 1/c - 1}{n} (1 + c\lambda)^{-1/c} \left(\frac{c\lambda}{1 + c\lambda}\right)^n (\phi_S(t))^n$$

$$\phi_F(t) = \sum_{n=0}^{\infty} \binom{n + 1/c - 1}{n} (1 + c\lambda)^{-1/c} \left(\frac{c\lambda\phi_S(t)}{1 + c\lambda}\right)^n$$

$$\phi_F(t) = \sum_{n=0}^{\infty} \binom{n + 1/c - 1}{n} (1 + c\lambda)^{-1/c} \left(\frac{cx}{1 + cx}\right)^n$$

where $x = \dfrac{\phi_S(t)}{1 - c\lambda(\phi_S(t) - 1)}$

$$\phi_F(t) = (1 + c\lambda)^{-1/c} (1 + cx)^{1/c}$$

$$\phi_F(t) = (1 - c\lambda(\phi_S(t) - 1))^{-1/c} \tag{5.16}$$

Note that Equations 5.14 and 5.16 are identical except for the different interpretation of the contagion parameter c. It should also be noted that the expression in Equations 5.14 and 5.16 approaches the expression in Equation 5.15 as c approaches 0.

In the computer program described below, we set $c = 10^{-7}$ whenever $|c| < 10^{-7}$. Thus the same computer code handles all three cases.

6. THE AGGREGATE LOSS DISTRIBUTION

In the preceding section we derived the characteristic function for the aggregate loss distribution for a single coverage or exposure class. In this section we use the above results to derive formulas for the cumulative probabilities and the excess pure premiums for multiple coverages or exposure classes.

For the sake of convenience, we make the following definitions.

$F(x)$ = Cumulative distribution function of the aggregate losses for all coverages combined

μ = Mean of aggregate loss distribution

σ = Standard deviation of aggregate loss distribution

$f(t)$ = modulus $(\phi_F(t/\sigma))$

$g(t)$ = argument $(\phi_F(t/\sigma))$

For each coverage, j, we define the following.

$$h_j(t) = \bar{h}_j(t/\sigma) - 1 \tag{6.1}$$

$$k_j(t) = \bar{k}_j(t/\sigma) \tag{6.2}$$

where $\bar{h}_j$ and $\bar{k}_j$ are given in Equations 5.12 and 5.13.

Note the $F(x)$ is the convolution of the aggregate loss distributions for each individual coverage. Using Equations 5.4, 5.5, 5.9 and 5.12–5.16 we have the following.

$$f(t) = \prod_j \text{modulus} \quad (1 - c_j\lambda_j(\phi_{S_j}(t/\sigma) - 1))^{-1/c_j}$$

$$f(t) = \prod_j \text{modulus} \quad (1 - c_j\lambda_j(h_j(t) + ik_j(t)))^{-1/c_j}$$

$$f(t) = \prod_j ((1 - c_j\lambda_j h_j(t))^2 + (c_j\lambda_j k_j(t))^2)^{-1/2c_j} \tag{6.3}$$

$$g(t) = \sum_j \text{argument} \quad (1 - c_j\lambda_j(\phi_{S_j}(t/\sigma) - 1))^{-1/c_j}$$

$$g(t) = \sum_j \text{argument} \quad (1 - c_j\lambda_j(h_j(t) + ik_j(t)))^{-1/c_j} \tag{6.4}$$

Once the modulus and the argument of the aggregate characteristic have been determined, it is possible to calculate the cumulative probabilities by use of the following formula.

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{f(t)}{t} \sin\,(tx/\sigma - g(t))\,dt \tag{6.5}$$

The excess pure premium can be obtained from the cumulative distribution function by the following formula.

$$EP(x) = \int_x^\infty (u - x)dF(u)$$

Applying this to Equation 6.5 we get the following formula.

$$EP(x) = \mu - \frac{x}{2} + \frac{\sigma}{\pi}\int_0^\infty \frac{f(t)}{t^2}\,(\cos\,(g(t)) - \cos\,(tx/\sigma - g(t)))dt \tag{6.6}$$

The excess pure premium ratio is defined by the following formula.

$$ER(x) = EP(x)/\mu$$

We now introduce parameter uncertainty of the severity distributions.

$$\mathcal{F}(x) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{f(t)}{t}\left(1 + \left(\frac{xt}{r\sigma}\right)^2\right)^{-(1+r)/2} \sin\left((1 + r)\arctan\left(\frac{xt}{r\sigma}\right) - g(t)\right) dt \tag{6.7}$$

$$\mathcal{EP}(x) = \mu - \frac{x}{2} + \frac{\sigma}{\pi}\int_0^\infty \frac{f(t)}{t^2}\left[\cos\,(g(t)) - \left(1 + \left(\frac{xt}{r\sigma}\right)^2\right)^{-r/2} \cos\left(r \cdot \arctan\left(\frac{xt}{r\sigma}\right) - g(t)\right)\right] dt \tag{6.8}$$

In the above two formulas, $r = 1 + 1/b$.

Equations 6.5–6.8 are derived in Appendix A. It should be noted that Equation 6.5 is the limit of Equation 6.7 as b approaches 0. Similarly Equation 6.6 is the limit of Equation 6.8 as b approaches 0. In our program we set $b = 10^{-7}$ whenever $b < 10^{-7}$ and thus the same computer code handles both situations.

Equations 6.7 and 6.8 are set up so that the parameter uncertainty of claim severity affects all coverages in the same way. This may be realistic if one believes that uncertainty in claim severity is due to inflation and that inflation affects all coverages in the same way. If one wants parameter uncertainty of claim severity for each coverage to be independent, several runs of the program will be required. An example showing how to do this will be given below.

7. NUMERICAL INTEGRATION

We now turn to the problem of evaluating the integrals given in Equations 6.7 and 6.8. It should also be noted that our program is written in FORTRAN to run on a large (IBM 370) computer. In this environment, it gets nearly instantaneous response at the computer terminal. The same algorithm has also been coded in BASIC to run on a TRS80 Model III microcomputer where it reproduces the mainframe results though with substantially greater running time. The actual FORTRAN code is included as Exhibit IX.

We now outline our algorithm. Explanation for the steps will be given below.

<u>Step</u>

1. Enter the parameters for the claim severity and the claim count distributions.
2. Calculate the aggregate mean, μ, and standard deviation, σ.
3. Enter the loss amounts, x.
4. Calculate basic interval length, h.
   $h = 2\pi\sigma$ /(maximum loss amount)
5. In order to apply the Gaussian quadrature formulas, we must evaluate the integrands at specified points. We evaluate the functions $f(t)$ and $g(t)$ at the appropriate points in each of the following intervals.

Interval Number	Interval
1	$(0, h/16)$
2	$(h/16, h/8)$
3	$(h/8, h/4)$
4	$(h/4, h/2)$
5	$(h/2, h)$
6	$(h, 2h)$
.	
.	
.	
$j + 4$	$((j - 1)h, jh)$

j is determined so that $f(t)t < .00002$ for all values of t evaluated in the interval $((j - 1)h, jh)$.

6. For each loss amount, x, evaluate $\mathcal{F}(x)$ and $\mathcal{EP}(x)$ by summing the results of the Gaussian quadrature formulas over each of the intervals given in Step 5.

We now give a more detailed explanation of the above steps.

Step 1

The parameters for each claim severity distribution are the claim severities $a_1, \ldots, a_{n+1}$ and the associated probabilities $p_1, \ldots, p_n$.

The parameters for each claim count distribution are the expected number of claims and the contagion parameter, c. Note that if $|c| < 10^{-7}$, we substitute $c = 10^{-7}$.

We must also enter the mixing parameter, b. If $b < 10^{-7}$ we substitute $b = 10^{-7}$.

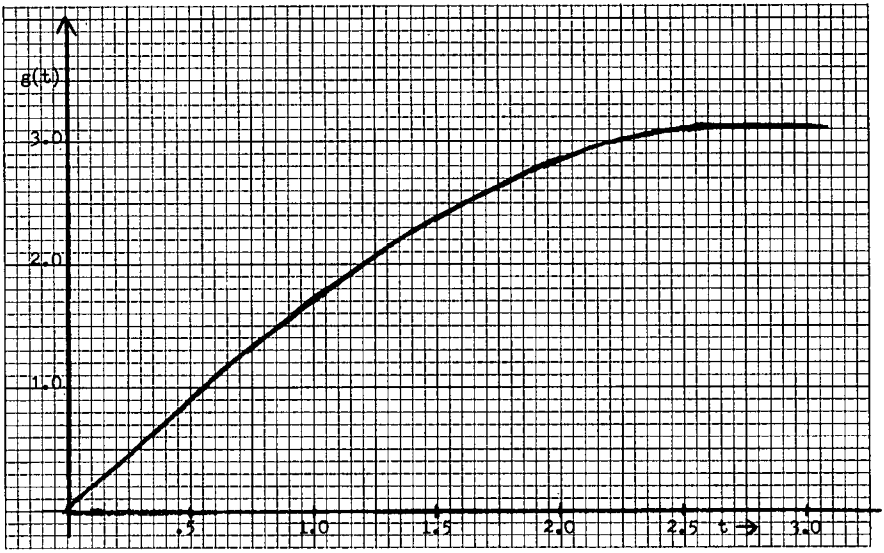
Step 2

For each coverage we calculate the aggregate mean and variance according to Equations 3.6 and 3.7. The aggregate mean and variance are the sums of the individual means and variances for each coverage.

Step 4

Evaluating a typical $g(t)$ showed that $g(t)$ changes slowly. See Figure 2. Also, $r \cdot \arctan(xt/r\sigma)$ is an increasing function of t which is bounded by xt/σ. Thus by choosing $h = 2\pi\sigma/(\text{maximum loss amount})$ we assure that the interval of integration will contain no more than one oscillation of the integrand.

FIGURE 2



Step 5

The evaluation of $f(t)$ and $g(t)$ is the most time consuming operation of this entire algorithm. Thus $f(t)$ and $g(t)$ should only be evaluated once for any given value of t, and the number of points, t, at which these functions are evaluated should be as few as possible. Inspection of the integrands of Equations 6.7 and 6.8 revealed that they changed most rapidly in the interval $(0, h)$. See Figures 3 and 4. Thus it was felt that the intervals used in the numerical integration should be relatively short in the interval $(0, h)$.

By a change of variables, each interval of integration was transformed from the given interval to the interval $(-1, 1)$. The Gaussian 5-point formula is then applied. The points, t_j, where $f(t)$ and $g(t)$ must be evaluated are as follows.

$$t_1 = (-0.90617985\ (b - a) + (b + a))\ /2$$
$$t_2 = (-0.53846931\ (b - a) + (b + a))\ /2$$
$$t_3 = (b + a)\ /2$$
$$t_4 = (0.53846931\ (b - a) + (b + a))\ /2$$
$$t_5 = (0.90617985\ (b - a) + (b + a))\ /2$$

Here a is the left endpoint of the interval, and b is the right endpoint of the interval under consideration. If $f(t_j)\ /\ t_j < .00002$ for $j = 1, \ldots, 5$ or the number of intervals equals 256, no more intervals are used.

FIGURE 3

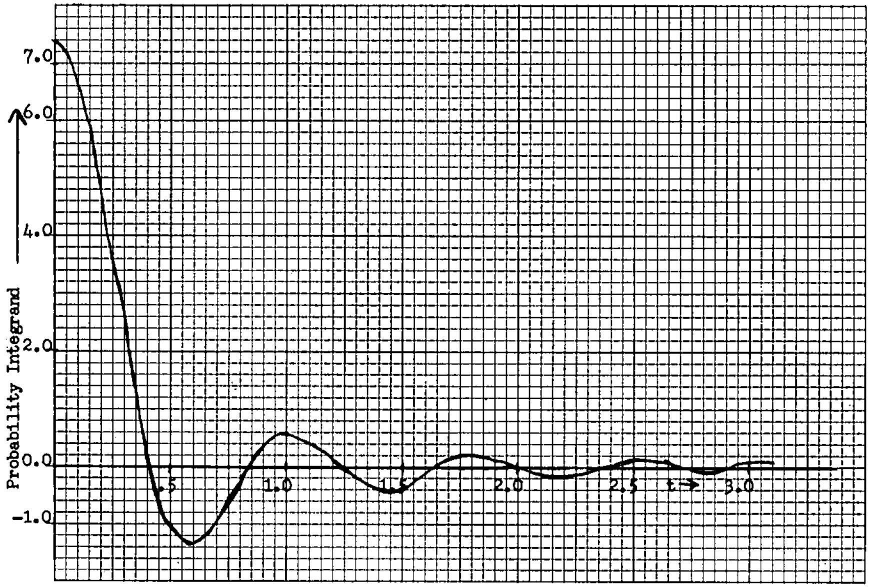
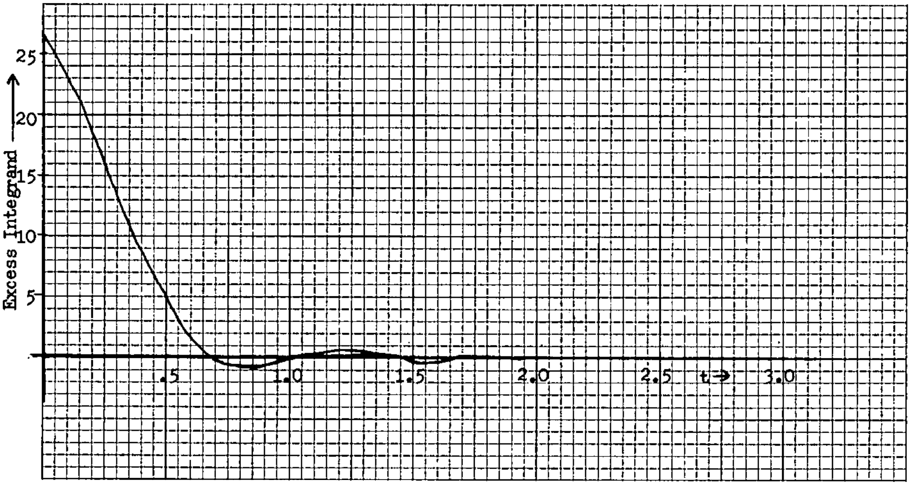


FIGURE 4



Step 6

Now that $f(t)$ and $g(t)$ are evaluated and stored in an array, it becomes an easy task to evaluate $\mathcal{F}(x)$ and $\mathcal{EP}(x)$. For each interval of integration we use the following rule to evaluate the integral.

$$I_p = \left(\sum_{j=1}^{5} W_j P\,(t_j)\right) \cdot \text{(interval length/2)}$$

$$I_E = \left(\sum_{j=1}^{5} W_j Q\,(t_j)\right) \cdot \text{(interval length/2)}$$

$$\text{where } P(t) = \frac{f(t)}{t}\left(1 + \left(\frac{xt}{r\sigma}\right)^2\right)^{-(1+r)/2} \sin\left((1 + r)\arctan\left(\frac{xt}{r\sigma}\right) - g(t)\right)$$

$$Q(t) = \frac{f(t)}{t^2}\left[(\cos\,(g(t)) - \left(1 + \left(\frac{xt}{r\sigma}\right)^2\right)^{-r/2} \cos\left(r \cdot \arctan\left(\frac{xt}{r\sigma}\right) - g(t)\right)\right]$$

$$W_1 = W_5 = 0.23692689$$
$$W_2 = W_4 = 0.47862867$$
$$W_3 = 0.56888889.$$

Then $\mathcal{F}(x) = .5 + (\text{Sum of all the } I_p\text{'s})\,/\pi$ and
$\mathcal{EP}(x) = \mu - x/2 + (\text{Sum of all the } I_E\text{'s})\,\sigma/\pi.$

8. ERROR ANALYSIS

There are three sources of error in the above calculations.

Roundoff Error

We use double precision arithmetic at every stage of our calculation. Double precision numbers are accurate to 16 significant digits on IBM equipment. Even though the calculations leading to a particular output value could number in the hundreds, it is doubtful that accumulated roundoff error could be an important factor in our calculations.

Discretization Error

The discretization error for the Gaussian 5-point formula is given by the expression

$$\frac{f^{(10)}(\xi)}{11 \cdot 7^3 \cdot 5^2 \cdot 3^8 \cdot 2} = 8.08 \times 10^{-10} \cdot f^{(10)}(\xi),\ \xi \text{ in } (-1,1).$$

Since the integrands are reasonably smooth (see Figures 3 and 4) the bound on $f^{(10)}$ should be reasonable. Thus the discretization error should not be significant.

Truncation Error

The most significant source of error in these calculations is the truncation error, or the error made by substituting an integral with finite limits of integration for an integral with infinite limits of integration. We now turn to analyzing this truncation error.

The truncation error, E_T, for the excess pure premium is given by

$$E_T = \frac{\sigma}{\pi} \int_a^\infty \frac{f(t)}{t^2} \left[\cos(g(t)) - \left(1 + \left(\frac{xt}{r\sigma}\right)^2\right)^{-r/2} \cos\left(r \cdot \arctan\left(\frac{xt}{r\sigma}\right) - g(t)\right)\right] dt$$

where a is the limit of the finite integral.

$$\text{Now } |E_T| \le \frac{\sigma}{\pi} \int_a^\infty \frac{f(t)}{t^2}(1 + 1)dt$$

$$\le \frac{2\sigma}{\pi} \cdot \max_{t \ge a}(f(t)) \int_a^\infty \frac{dt}{t^2}$$

$$= \frac{2\sigma}{\pi} \cdot \max_{t \ge a}(f(t)) \cdot \frac{1}{a}. \tag{8.1}$$

$$\text{Now } f(t) = \left| \int_0^\infty e^{itx} dF(x) \right| \le \int_0^\infty |e^{itx}|\, dF(x) = 1. \tag{8.2}$$

$$\text{Thus } E_T \le \frac{2\sigma}{\pi} \cdot \frac{1}{a}.$$

The bound on the truncation error given by Equation 8.2 is extremely conservative because, as we show in Appendix B, $\max_{t \geq a} f(t)$ will be significantly less than one for most cases of interest. In fact, if each (piecewise linear) claim severity distribution function is continuous, $f(t)$ approaches the probability of zero claims as t approaches infinity. For example, when the claim count distribution is Poisson with a mean of 10 claims, $f(t)$ will be close to e^{-10} or 0.0000454 for large t.

The bound on the truncation error given by Equation 8.1 is also conservative because the integrand repeatedly changes sign.

In our program, a is usually chosen so that $\max_{t \geq a} (f(t)) \cdot 1/a < .00002$. Thus we would expect the truncation error for the excess pure premium ratio to be bounded by $.000013 \cdot \sigma/\mu$.

The truncation error for the cumulative probabilities does not permit an analysis similar to the above because the denominator of the integrand contains t instead of t^2. The examples in the next section will show that cumulative probabilities calculated by this algorithm seem to be accurate. But they are somewhat less accurate than the excess pure premium.

9. NUMERICAL TESTS OF THE ALGORITHM

There are cases when the algorithm can be compared with known results. We consider two such cases.

If the contagion parameter, c, is equal to -1, then $\phi_F(x) = \phi_S(x)$. The choice of $c = -1$ corresponds to the Binomial distribution with $m = p = 1$.

For our first example, consider the following.

$$F(x) = S(x) = x \text{ for } 0 \leq x \leq 1$$

$$ER(x) = \frac{1}{\mu} \int_x^1 (1 - F(u))du = (1 - x)^2$$

Table 9.1 compares computed to actual results.

TABLE 9.1

x	$F(x)$ Actual	$F(x)$ Computed	$ER(x)$ Actual	$ER(x)$ Computed
.10	.1000	.1000	.8100	.8100
.20	.2000	.2000	.6400	.6400
.30	.3000	.3000	.4900	.4900
.40	.4000	.4000	.3600	.3600
.50	.5000	.5000	.2500	.2500
.60	.6000	.6000	.1600	.1600
.70	.7000	.7000	.0900	.0900
.80	.8000	.8000	.0400	.0400
.90	.9000	.9000	.0100	.0100
1.00	1.0000	.9995	.0000	.0000

For our next example, consider the following.

$F(x) = S(x) = x/2$ for $0 \leq x < 1$
$F(1) = S(1) = 1$

$$ER(x) = \frac{1}{\mu}\int_x^1 (1 - F(u))du = (3 - x)(1 - x)/3$$

For reasons described in Section 7 above, the program required 256 intervals for the numerical integration. The value of $f(t)/t$ for the largest value of t was equal to .001. Using Equation 8.2 we obtained an estimate of .00027 as a bound on the truncation error for $ER(x)$. Table 9.2 compares computed to actual results.

These examples would seem to indicate that the calculation of $ER(x)$ is more accurate than that of $F(x)$. If $F(x)$ is continuous, the error appears to be small, but, if $F(x)$ is not continuous, the errors may not be so small near the points of discontinuity.

We now turn to a more realistic example. Exhibit II shows an actual run of our program. Details concerning the input will be given in the discussion of aggregate increased limits factors which follows. Here we provide a comparison between the results of our program and a Monte Carlo simulation. One should not expect exact agreement between expected and observed results due to

TABLE 9.2

x	$F(x)$ Actual	$F(x)$ Computed	$ER(x)$ Actual	$ER(x)$ Computed
.10	.0500	.0501	.8700	.8700
.20	.1000	.1001	.7467	.7467
.30	.1500	.1502	.6300	.6300
.40	.2000	.2002	.5200	.5200
.50	.2500	.2502	.4167	.4167
.60	.3000	.3003	.3200	.3200
.70	.3500	.3504	.2300	.2300
.80	.4000	.4005	.1467	.1467
.90	.4500	.4511	.0700	.0700
.99	.4950	.4869	.0067	.0067
1.00	1.0000	.7499	.0000	.0001
1.01	1.0000	1.0081	.0000	.0000
1.05	1.0000	.9979	.0000	.0000

simulation error. For this reason we performed a Chi-Square test on the results to see if the difference could be explained by random fluctuations. The results are in Table 9.3.

The expected number of claims in each cell was obtained from Exhibit II. The observed number of claims in each cell was obtained by a Monte Carlo simulation using exactly the same input parameters as those in Exhibit II. Ten thousand trials were used.

If the differences between observed and expected values are due solely to random fluctuations, one should expect a Chi-Square value of 25. In this case we get a slightly higher value of Chi-Square. We have performed similar tests on many occasions and have gotten similar results. The algorithm works.

10. AGGREGATE LIMITS

We now consider how this algorithm can be used to calculate the premium for a policy that is subject to an aggregate limit.

Underwriters have long felt that lines of insurance such as Products Liability and Medical Malpractice present a severe catastrophe potential. For example,

TABLE 9.3

CHI-SQUARE TEST FOR AGGREGATE LOSS DISTRIBUTIONS

Upper Cell Boundary	Observed	Expected
50,000	51	52
100,000	273	268
150,000	432	435
200,000	546	540
250,000	589	587
300,000	632	628
350,000	736	737
400,000	782	782
450,000	789	769
500,000	721	720
550,000	641	662
600,000	625	622
650,000	597	561
700,000	506	491
750,000	402	416
800,000	353	349
850,000	269	294
900,000	227	241
950,000	201	195
1,000,000	135	154
1,050,000	93	121
1,100,000	102	94
1,150,000	73	73
1,200,000	46	55
1,250,000	39	42
Over 1,250,000	140	112

Chi-Square = 26.0
Degrees of Freedom = 25

the publicity given a Products Liability lawsuit may well provoke several additional lawsuits by others who have purchased the same product. Thus underwriters have justifiably sought to limit the total amount of losses that can be paid out under a single policy.

The price for a policy with an aggregate limit (ignoring expense considerations) will be the price of a similar policy with no aggregate limit less the excess pure premium for the aggregate limit. Below, we will give several examples of such calculations using Exhibits II to V. But, before we do this, let us discuss the input parameters.

The claim severity distribution chosen is typical for Products Liability coverages. We will not discuss selection of the claim severity distribution here. Instead we will refer the interested reader to the literature [19] [20].

The claim severity distribution will be subject to a $250,000 occurrence limit.

The mean of the claim count distribution was calculated by dividing total expected losses by the severity mean ($18,198). In Exhibits II, IV and V a contagion parameter of zero was chosen. This choice gives the Poisson distribution. In Exhibit III we chose a contagion parameter of .25. In light of the catastrophe potential for Products Liability that was discussed above, a more highly skewed claim count distribution would indeed seem justified.

A mixing parameter of 0 is used in this example.

Tables 10.1 and 10.2 show the discounts expressed as a proportion of the total expected loss.

While a more highly skewed claim count distribution may be justified for Products Liability, it does not give a conservative price for a policy with an aggregate limit. Thus we would recommend using a Poisson distribution for the claim count unless one has definite evidence that a more skewed distribution is appropriate.

Notice that the discounts depend upon the expected loss. Present tables of increased limits factors do not reflect this dependence. We admit that there is a practical problem involved in publishing increased limits factors that vary by expected loss. However, the "practical" solution of not considering the expected loss can produce embarrassing examples such as the following. This method is identical to that given in I.S.O. rating manuals.

TABLE 10.1

DISCOUNTS FOR AGGREGATE LIMITS

Expected Loss = $500,000

	Contagion Parameter	
Aggregate Limit	0.00	0.25
$ 600,000	.1394	.2132
800,000	.0516	.1125
1,000,000	.0165	.0570
1,200,000	.0046	.0279
1,400,000	.0012	.0133

TABLE 10.2

DISCOUNTS FOR AGGREGATE LIMITS

Contagion Parameter = 0.0

	Expected Loss		
Aggregate Limit	$250,000	$500,000	$1,000,000
$ 600,000	.0296	.1394	.4202
$ 800,000	.0060	.0516	.2665
$1,000,000	.0010	.0165	.1528
$1,200,000	.0002	.0046	.0791
$1,400,000	—	.0012	.0371

Basic Limits — $25,000 per occurrence and $75,000 aggregate
Base Rate — $1.00 per unit of exposure
Exposure — 1,000,000 units

If an insured bought a policy for the basic limits, he would pay $1,000,000 and the most he could recover in losses is $75,000! While it is unlikely that such a policy has ever been sold, significant errors could be quite common.

We propose the following as a remedy to this situation.

1. Publish increased limits tables for occurrence limits only.
2. Do not give discounts for aggregate limits. Instead, publish a table of aggregate limits which are appropriate for a given expected loss. The aggregate limits should be sufficiently high so that the indicated discount is less than a nominal amount, say 0.5%.

Using Exhibits II, IV and V we can derive the appropriate aggregate limits.

Expected Loss	Aggregate Limit
$ 250,000	$ 825,000
500,000	1,200,000
1,000,000	1,900,000

11. GROUP LIFE AGGREGATE EXCESS INSURANCE

We now give the solution to a problem that was proposed to us by a life actuary of our company.

A large employer wanted to self insure his group life insurance. To protect against a catastrophe, he wanted to purchase aggregate excess insurance to cover losses in excess of 1.25 times the expected loss. The following data were provided to us.

Group	Age Range	Number of Lives	Expected Loss
1	29 and Under	2,073	47,086
2	30–34	1,135	36,342
3	35–39	1,044	35,380
4	40–44	822	54,938
5	45–49	1,004	136,126
6	50–54	1,193	270,050
7	55–59	975	395,471
8	60–64	546	258,525
9	65 and Over	25	13,247

The expected loss was computed using a mortality table and the average amount of insurance in each group.

It was felt that the claim count distribution should be binomial. Thus we chose a contagion parameter of $-1/(\text{number of lives})$ for each group. We were not given a distribution of insurance amounts for each group. Assuming that all insureds had the average amount of insurance in each group would understate the excess pure premium. For this reason we requested rough estimates for those distributions.

The mixing parameter selected was 0.0.

It should be noted that the assumptions of the collective risk model are violated in this example. When a person dies, the amount of his insurance policy is removed from the claim severity distribution. However the turnover of group members should keep the claim severity distribution approximately the same. Thus we feel that the collective risk model will be a good approximation of the true situation.

Exhibit VI gives the computer run for the problem. The pure premium for this coverage was calculated to be 1.53% of the expected loss.

12. RETROSPECTIVE RATING; NESTED AGGREGATES

A retrospective rating plan is a rating plan in which the final premium is determined after the policy period has expired [21]. While these plans have many features, we will limit this discussion to plans where the insurer is liable for all losses above an agreed upon amount.

Retrospective rating plans can cover several different policies under a single plan. Here we provide a simple example showing how to calculate the pure premium, or insurance charge, for such a rating plan. Our example will consist of two coverages, Workers' Compensation and Products Liability.

The Workers' Compensation policy has an expected loss of $500,000. The claim severity distribution is given in Exhibit I. The contagion parameter is .05. The mixing parameter is 0.0.

The Products Liability policy has an expected loss of $500,000 before application of the aggregate limit. The claim severity distribution is given in Exhibit I. The contagion parameter is .25. The policy that is written under the retrospective rating plan is subject to a $1,000,000 aggregate limit. The mixing parameter is 0.0.

The presence of a policy subject to an aggregate limit in the retrospective rating plan makes it necessary to run the program twice to determine the insurance charges. Exhibit III will serve as the first run of the program. For the second run we treat the Workers' Compensation parameters in the usual manner. For the Products Liability, we substitute the aggregate loss distribution in Exhibit

III for the claim severity distribution. We, of course, limit the aggregate losses to $1,000,000. The contagion parameter is -1. This corresponds to a binomial claim count distribution with $m = p = 1$. The results of the second run are shown in Exhibit VII. It can be seen, for example, that the insurance charge for a plan which covers losses in excess of $1,500,000 is $21,894.

We now consider parameter uncertainty for the scale of the claim severity distribution.

Misestimation of the claim severity distribution can occur because of limited information on the individual coverage. In this case one would expect the scale uncertainty for each coverage to be independent. Misestimation of future inflation can also cause scale uncertainty. In this case one could expect the scale uncertainty to affect each coverage in the same way. The following example shows how to handle both of these cases. It will be necessary to run the program once for each individual coverage. A final run is then required to combine the individual coverages.

The Workers' Compensation policy has the same parameters that were specified in the above example with the exception that the mixing parameter is set equal to .05. This reflects uncertainty in the scale of the claim severity distribution for Workers' Compensation. The aggregate loss distribution for this coverage is given in Exhibit VIIIA.

The Products Liability policy has the same parameters that were specified in the above example with the exception that the mixing parameter is set equal to .05. This reflects uncertainty in the scale of the claim severity distribution for Products Liability. The aggregate loss distribution is given in Exhibit VIIIB. It should be noted that this aggregate loss model adjusts the policy limit with the scaling parameter, while in the real world the policy limit remains fixed. However this should not significantly affect the final results.

The aggregate loss distributions for Workers' Compensation and Products Liability are then combined to get the aggregate loss distribution for the total losses of the two coverages. Here the aggregate loss distribution for each coverage is treated as the claim severity distribution for the final run of the program. The Products Liability loss is limited to $1,000,000. The contagion parameter for each coverage is set equal to -1. The mixing parameter is set equal to .05. This reflects uncertainty in the scale of the aggregate loss distribution. For a given year the scale parameter is identical for both coverages. It should be noted that, as we do above, this aggregate loss model adjusts the aggregate limit with the scaling factor.

The results of the third run are shown in Exhibit VIIIC. It can be seen, for example, that the insurance charge for a plan which covers losses in excess of $1,500,000 is $46,424. Here one can see that that parameter uncertainty can significantly affect the required insurance charge.

13. CONCLUSION

We have described an efficient and accurate algorithm which calculates the cumulative probabilities and excess pure premiums for the collective risk model.

The program and related programs have been used at our company in applications described above and many others. These include the analysis of profit sharing plans, large account pricing, aggregate deductibles and sliding scale dividend plans. Also, we are currently exploring applications involving the optimization of reinsurance retentions [15] and designing a retrospective rating plan which properly accounts for the "overlap" problem [22]. In short, this is a very useful program.

Our exposure to these applications has led us to believe that further work needs to be done with the collective risk model. In particular, we need to test the predictions of the collective risk model against actual aggregate loss data. We also need to test the sensitivity of the collective risk model to violations of the assumptions underlying it.

14. ACKNOWLEDGEMENTS

This paper had its origins in an anlysis of Shaw Mong's paper [10] which was done by Glenn Meyers and Nathaniel Schenker. Comparisons of Mong's results with Monte Carlo simulations suggested that Mong's technique worked well when the claim severity distribution was a Gamma distribution, but otherwise it worked poorly. It was also noted that Mong's technique could be modified to work for any claim severity distribution provided one could calculate its characteristic function.

The classic book on risk theory by Beard, Pentikäinen and Pesonen [7] had a very strong influence in our thinking, as can be noted by our several references to it. We regard this paper, in part, as a synthesis of the ideas in Mong's paper and Chapter 8 of the book.

Another strong influence has been observing the various ways this algorithm, and prior Monte Carlo simulations, have been used. We received excellent feedback from the following individuals: Burton Covitz, Michael Larsen, John Meeks, Arthur Placek, and Philip Wolf. The following individuals made several helpful comments while we were preparing this paper: Bradley Alpert, Yakov Avichai, Sam Gutterman, Phillip Norton, Nathaniel Schenker, and Edward Seligman. We offer our sincere thanks.

APPENDIX A
DERIVATION OF EQUATIONS 6.5–6.8

The purpose of this appendix is to derive Equations 6.5–6.8. We will first derive expressions for the cumulative probability and the excess pure premium in terms of $\phi_F(t)$ and $U(\beta)$. Equations 6.5–6.8 will then be special cases of these expressions.

The following formula is given by Kendall and Stuart [23].

$$F(\beta x) = \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty} \frac{e^{i\beta xt}\cdot\phi_F(-t) - e^{-i\beta xt}\cdot\phi_F(t)}{it}\,dt$$

From Equation 3.5 we have the following.

$$\begin{aligned}
\mathcal{F}(x) &= \int_0^{\infty} F(\beta x)dU(\beta) \\
&= \int_0^{\infty}\left[\frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty}\frac{e^{i\beta xt}\,\phi_F(-t) - e^{-i\beta xt}\,\phi_F(t)}{it}dt\right]dU(\beta) \\
&= \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty}\frac{1}{it}\left[\phi_F(-t)\int_0^{\infty} e^{i\beta xt}dU(\beta) - \phi_F(t)\int_0^{\infty} e^{-i\beta xt}dU(\beta)\right]dt \\
&= \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty}\frac{1}{it}\left[\phi_F(-t)\cdot\phi_U(xt) - \phi_F(t)\cdot\phi_U(-xt)\right]dt \qquad (A.1)
\end{aligned}$$

Thus we have the following.

$$\begin{aligned}
\mathcal{E}\mathcal{P}(x) &= \int_x^{\infty}(v - x)d\mathcal{F}(v) = \int_x^{\infty}\left[\int_x^{v} du\right]d\mathcal{F}(v) \\
&= \int_x^{\infty}\left[\int_u^{\infty} d\mathcal{F}(v)\right]du = \int_x^{\infty}(1 - \mathcal{F}(u))du \\
&= \int_0^{\infty}(1 - \mathcal{F}(v))dv - \int_0^{x}(1 - \mathcal{F}(v))dv
\end{aligned}$$

$$= \mu - \int_0^x (1 - \mathcal{F}(v))dv$$

$$= \mu - \frac{x}{2} - \int_0^\infty \frac{1}{2\pi}\left\{\int_0^\infty \frac{\phi_F(-t)\cdot\phi_U(vt) - \phi_F(t)\cdot\phi_U(-vt)}{it}\,dt\right\}dv$$

$$= \mu - \frac{x}{2} + \frac{1}{2\pi}\int_0^\infty \frac{1}{it}\left[\phi_F(-t)\int_0^x \phi_U(vt)dv - \right.$$

$$\left. \phi_F(t)\int_0^x \phi_U(-vt)dv\right]dt \qquad \text{(A.2)}$$

Note: $\phi(t) = f(t)e^{ig(t)}$; $f(t) = f(-t)$ and $g(-t) = -g(t)$.

Case 1 $U(\beta) = 0$ for $\beta < 1$ and $U(\beta) = 1$ for $\beta \geq 1$.

$$\phi_U(t) = e^{it}$$

$$\phi_U(xt) = e^{ixt} \qquad \text{(A.3)}$$

$$\phi_U(-xt) = e^{-ixt} \qquad \text{(A.4)}$$

$$\int_0^x \phi_U(vt)dv = \frac{e^{ixt} - 1}{it} \qquad \text{(A.5)}$$

$$\int_0^x \phi_U(-vt)dv = \frac{1 - e^{-ixt}}{it} \qquad \text{(A.6)}$$

Equation 6.5 is obtained by substituting Equations A.3 and A.4 into Equation A.1. and replacing t with t/σ. Equation 6.6 is obtained by substituting Equations A.5 and A.6 into A.2 and replacing t with t/σ.

Case 2 $dU(\beta) = \frac{r}{\Gamma(r+1)}(r\beta)^r e^{-r\beta}d\beta$

We first show that $U(\beta)$ satisfies the conditions stated for Algorithm 3.3.

$$E(1/\beta) = \frac{r}{\Gamma(r+1)}\int_0^\infty \frac{1}{\beta}(r\beta)^r e^{-r\beta}d\beta$$

$$= \frac{r}{\Gamma(r)}\int_0^\infty (r\beta)^{r-1}e^{-r\beta}d\beta$$

$$= 1$$

$$E(1/\beta^2) = \frac{r}{\Gamma(r+1)} \int_0^\infty \frac{1}{\beta^2} (r\beta)^r e^{-r\beta} d\beta$$

$$= \frac{r}{r-1} \frac{r}{\Gamma(r-1)} \int_0^\infty (r\beta)^{r-2} e^{-r\beta} d\beta$$

$$= \frac{r}{r-1}$$

If $r = 1 + 1/b$ we have that $\mathrm{Var}\,(1/\beta) = (r/(r-1)) - 1 = b$.

$$\phi_U(t) = \left(1 - \frac{it}{r}\right)^{-(1+r)}$$

$$\phi_U(xt) = \left(1 - \frac{ixt}{r}\right)^{-(1+r)} = \left(1 + \left(\frac{xt}{r}\right)^2\right)^{-(1+r)/2} e^{i(1+r)\arctan(xt/r)} \tag{A.7}$$

$$\phi_U(-xt) = \left(1 + \frac{ixt}{r}\right)^{-(1+r)} = \left(1 + \left(\frac{xt}{r}\right)^2\right)^{-(1+r)/2} e^{-i(1+r)\arctan(xt/r)} \tag{A.8}$$

$$\int_0^x \phi_U(vt)dv = \frac{1}{it}\left[\left(1 - \frac{ixt}{r}\right)^{-r} - 1\right]$$

$$= \frac{1}{it}\left[\left(1 - \left(\frac{xt}{r}\right)^2\right)^{-r/2} e^{ir\arctan(xt/r)} - 1\right] \tag{A.9}$$

$$\int_0^x \phi_U(-vt)dv = \frac{1}{it}\left[1 - \left(1 + \frac{ixt}{r}\right)^{-r}\right]$$

$$= \frac{1}{it}\left[1 - \left(1 + \left(\frac{xt}{r}\right)^2\right)^{-r/2} e^{-ir\arctan(xt/r)}\right] \tag{A.10}$$

Equation 6.7 is obtained by substituting Equations A.7 and A.8 into Equation A.1 and replacing t with t/σ. Equation 6.8 is obtained by substituting Equations A.9 and A.10 into A.2 and replacing t with t/σ.

APPENDIX B
ASYMPTOTIC BEHAVIOR OF $f(t)$

In the error analysis of Section 8 we indicated that $\max_{t \geq a} f(t)$ could be significantly less than one for large a. We now give a demonstration of this fact. It will be sufficient to consider a single coverage or class of business.

We will adopt the following notation for use in this appendix.

$$D = 1 - \sum_{k-1}^{n} P_k$$

$$A = a_{n+1}$$

As $t \to \infty$, we have the following.

$$h(t) \to D \cos (At)$$
$$k(t) \to D \sin (At)$$
$$\phi_S(t) \to D(\cos (At) + i \sin (At))$$

Case 1 Binomial Distribution

$$\phi_F(t) = [1 + p\,(\phi_S(t) - 1)]^m$$

As $t \to \infty$, we have the following.

$$f(t) \to [(1 - p + D\,p \cos (At))^2 + (D\,p \sin (At))^2]^{m/2}$$
$$f(t) \to [(1 - p)^2 + 2\,D\,p \cos (At) + (D\,p)^2]^{m/2}$$

If $D = 0, f(t) \to (1 - p)^m$ which is equal to the probability of having no claims. If $D > 0$, $f(t)$ does not approach a limit, but the asymptotic upper bound of $f(t)$ can be obtained by setting $\cos (At) = 1$.

$$\max_{t \geq a} f(t) \to [(1 - p)^2 + 2\,D\,p + (D\,p)^2]^{m/2}$$

As an example, consider the case when $m = 100$, $p = .1$ and $D = .1$:

$$\max_{t \geq a} f(t) \to .0000905.$$

Case 2 Poisson Distribution

$$\phi_F(t) = e^{\lambda(\phi_S(t)-1)}$$

As $t \to \infty$, $f(t) \to e^{-\lambda} \cdot e^{\lambda D \cos(At)}$. If $D = 0$, $f(t) \to e^{-\lambda}$, which is equal to the probability of having no claims. If $D > 0$, $f(t)$ does not approach a limit, but the asymptotic upper bound of $f(t)$ can be obtained by setting $\cos(At) = 1$.

$$\max_{t \geq a} f(t) \to e^{-\lambda(1-D)}$$

As an example, consider the case when $\lambda = 10$ and $D = .1$:

$$\max_{t \geq a} f(t) \to e^{-9} = .0001234.$$

Case 3 Negative Binomial Distribution

$$\phi_F(t) = [1 - c\lambda(\phi_S(t) - 1)]^{-1/c}$$

As $t \to \infty$, we have the following.

$$f(t) \to [(1 + c\lambda + c\lambda\, D \cos(At))^2 + (c\lambda\, D \sin(At))^2]^{-1/2c}$$
$$f(t) \to [(1 + c\lambda)^2 + 2(1 + c\lambda) \cdot c\lambda\, D \cos(At) + (c\lambda\, D)^2]^{-1/2c}$$

If $D = 0$, $f(t) \to (1 + c\lambda)^{-1/c}$, which is the probability of having no claims. If $D > 0$, $f(t)$ does not approach a limit, but the asymptotic upper bound of $f(t)$ can be obtained by setting $\cos(At) = 1$.

$$\max_{t \geq a} f(t) \to [(1 + c\lambda)^2 - 2(1 + c\lambda)(c\lambda D) + (c\lambda D)^2]^{-1/2c}$$

As an example, consider the case when $\lambda = 10$, $D = .1$ and $c = .1$:

$$\max_{t \geq a} f(t) \to .001631.$$

NOTES AND REFERENCES

[1] Paul Dorweiler, "On Graduating Excess Pure Premium Ratios," *PCAS* XXVIII, 1941, p. 132.

[2] Nels M. Valerius, "Risk Distributions Underlying Insurance Charges in the Retrospective Rating Plan," *PCAS* XXIX, 1942, p. 96.

[3] LeRoy J. Simon, "The 1965 Table M," *PCAS* LII, 1965, p. 1.

[4] Charles C. Hewitt, Jr., "Loss Ratio Distributions— A Model," *PCAS* LIV, 1967, p. 70.

[5] Charles A. Hachemeister, "A Stochastic Model For Loss Reserving," *Transactions of the 21st International Congress of Actuaries,* Vol. 1980, p. 185.

[6] Insurance Services Office, *Report of the Increased Limits Subcommittee: A Review of Increased Limits Ratemaking,* Insurance Services Office, 1980, p. 72.

[7] R. E. Beard, T. Pentikäinen, and E. Pesonen, *Risk Theory,* 2nd Edition, Chapman and Hall, 1977, p. 43.

[8] Gary Venter, "Transformed Beta and Gamma Distributions and Aggregate Losses," *Pricing, Underwriting, and Managing the Large Risk,* Casualty Actuarial Society 1982 Discussion Paper Program, p. 395. See also *PCAS* LXX, (1983).

[9] Glenn Meyers, "An Analysis of Retrospective Rating," *PCAS* LXVII, 1980, p. 110.

[10] Shaw Mong, "Estimating Aggregate Loss Probability and Increased Limit Factor," *Pricing Property and Casualty Insurance Products,* Casualty Actuarial Society 1980 Discussion Paper Program, p. 358.

[11] Ethan Stroh, "Actuarial Note: The Distribution Functions of Collective Risk Theory as Linear Compounds," *Actuarial Research Clearing House,* 1978.1.

[12] James Tilley, Discussion of Harry Panjer, "The Aggregate Claim Distribution and Stop-Loss Reinsurance," *TSA* XXXII, 1980, p. 538.

[13] Harry Panjer, "Recursive Evaluation of a Family of Compound Distributions," *ASTIN Bulletin,* Vol. 12, No. 1, 1981, p. 22.

[14] Glenn Meyers and Nathaniel Schenker, "Parameter Uncertainty in the Collective Risk Model," *Pricing, Underwriting, and Managing the Large Risk,* Casualty Actuarial Society 1982 Discussion Paper Program, p. 253. See also *PCAS* LXX, (1983).

[15] Philip Heckman and Phillip Norton, "Optimization of Excess Portfolios," *Pricing, Underwriting, and Managing the Large Risk,* Casualty Actuarial Society 1982 Discussion Paper Program, p. 113.

[16] Beard *et al., op. cit.,* p. 18.

[17] *Ibid.,* p. 110.

[18] *Ibid.,* p. 126.

[19] Insurance Services Office, *op. cit.,* p. 46.

[20] Gary Patrik, "Estimating Casualty Insurance Loss Amount Distributions," *PCAS* LXVII, 1980, p. 57.

[21] National Council on Compensation Insurance, *Retrospective Rating Plan D.*

[22] For a discussion of the overlap problem, see Meyers *op. cit.*

[23] Maurice Kendall and Alan Stuart, *The Advanced Theory of Statistics,* Vol. I, 4th Edition, Macmillan, 1977, p. 98.

NOTE

The exhibits associated with the paper "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions" by Philip E. Heckman and Glenn G. Meyers (*PCAS* LXX, 1983) appear in the subsequent volume of the *Proceedings* (*PCAS* LXXI, 1984).