

[D7]  
**PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY  
FROM INDIVIDUAL PAYMENTS DATA**  
Contributed by T S Wright

**1. Introduction**

Many stochastic claims reserving methods use aggregate data and yield the first two moments (best estimate and standard error) of the outstanding liability. The term "aggregate data" here refers to the total of payments made for each cohort and each stage of development as opposed to the amount of each individual payment. This paper describes an approach which differs from most stochastic reserving methods in both these respects:

- (i) it makes use of data on the amount of each individual payment (rather than aggregate data)
- (ii) it yields a complete probability distribution for the total outstanding liability (rather than just the first two moments).

The advantage of point (ii) is that more complete information is provided for setting the claims reserve. For most purposes, a claims reserve should include an allowance for possible adverse experience rather than being purely the "best estimate" (that is, the first moment, or expected value). This extra allowance is analogous to a safety load (or risk load) in (re) insurance pricing, and we shall use this terminology.

The standard error is a fairly crude measure of uncertainty on which to base the safety load. It makes no allowance for possible skewness in the distribution of outstanding liabilities. Typically, the number of IBNR claims is best represented using a skew distribution such as the Negative Binomial. This is partly because there is usually some potential for a single cause to give rise to many more claims than expected, for example, a legal decision which sets a precedent. This in turn means that the total of all outstanding liabilities has a skew distribution.

The chance that the total of outstanding liabilities will exceed the best estimate plus some fixed multiple (e.g. two) of the standard error depends on the skewness. If this is relatively high, this fact should be reflected in the safety load: a stochastic method which yields only the first two moments does not provide the opportunity to do this. By calculating a complete probability distribution for the total outstanding liability, we have the ability to apply any safety load principle in setting the reserve. In particular, a safety load principle which takes account of all moments of the distribution can be applied, for example, the "proportional hazards" principle, Wang (1995).

Although it is possible to calculate a complete distribution of outstanding liabilities from aggregate paid claims data, individual payments data provide more information for this purpose, with a consequent increase in the reliability of the results (that is, a smaller standard error). It is a well known principle in statistics that aggregating data results in a loss of information: this is why the use of individual payments rather than aggregate data, as mentioned at (i) above, offers potential benefits.

There are exceptions to this principle: "sufficient statistics" are aggregations of data which do not involve any loss of information. However, specific statistics are "sufficient" only under specific stochastic assumptions. These assumptions might not be appropriate, and in any case, can be more thoroughly verified using the more detailed, pre-aggregated data.

## 2. Overview

There are four main stages to the approach described in this paper:

- (i) construct a probability distribution for the size of individual future claim payments
- (ii) construct a probability distribution for the number of future claim payments
- (iii) carry out a compounding calculation to combine the distributions calculated in steps (i) and (ii), in order to arrive at the probability distribution for the total of all future payments
- (iv) apply an appropriate safety load principle to determine the reserve from the probability distribution calculated at step (iii).

This paper focuses on steps (i) and (ii) because steps (iii) and (iv) are well defined problems on which there is a substantial literature. For step (iii), the calculation of a compound distribution, practical solutions are provided by Panjer (1981) and by Heckman & Meyers (1983). For step (iv), the calculation of a safety-loaded premium (or in the present context a claims reserve) from a loss distribution, a practical and theoretically consistent solution has been provided by Wang (1995). Although Wang's solution is not unduly complex, there are simpler approaches to step (iv) which are discussed briefly in section 7 of this paper. (These simpler approaches do have theoretical weaknesses, which could manifest themselves as inconsistencies in the allocation of the total reserve across different cohorts.)

## 3. Example Data

Each remaining section of the paper is illustrated by an example which is carried through all sections. This example is artificially simplified: we have only three accident years, with an unrealistically small number of claim payments in each one. Nevertheless, the example illustrates the main features of this approach.

PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY  
FROM INDIVIDUAL PAYMENTS DATA

The triangle below gives the number of individual claim payments:

*Number of claim payments*

		Development year		
		1	2	3
Accident year	1	10	19	5
	2	12	22	
	3	11		

The amount of each individual claim payment is given below. These amounts have been pre-adjusted for inflation at 5% by applying the factor 1.1025 to each of the 10 payments in the first diagonal, and the factor 1.05 to each of the 31 payments in the second diagonal of the run-off triangle. The adjusted amount of each payment is given in the first column below. The second column indicates the development year. The payments have been sorted into order of ascending size:

468.0	1	652.0	1	720.0	1	834.0	1	1018.0	1	1094.0	1
1268.0	2	1379.0	1	1408.0	1	1551.0	1	1620.0	1	1761.0	1
1870.0	1	1933.0	1	2022.0	1	2025.0	1	2158.0	1	2168.0	1
2225.0	2	2323.0	2	2462.0	1	2575.0	1	2579.0	1	2605.0	1
2676.0	1	2765.0	1	2820.0	1	2937.0	1	3135.0	2	3460.0	1
3502.0	1	3585.0	1	3639.0	2	3815.0	2	3938.0	1	3953.0	2
4187.0	2	4196.0	1	4259.0	1	4792.0	2	4960.0	2	5212.0	2
5253.0	2	5395.0	2	5865.0	1	5953.0	2	6094.0	2	6109.0	1
6210.0	2	6251.0	2	6319.0	2	6401.0	2	6615.0	2	6801.0	2
6905.0	2	7048.0	2	7062.0	2	7418.0	2	7589.0	2	7647.0	2
7886.0	2	8573.0	2	8661.0	2	9244.0	2	9515.0	2	9553.0	3
10043	2	10160	2	11826	2	11905	2	12968	2	13265	2
13609	3	16659	2	16756	2	18455	3	19875	2	22278	3
34091	3										

#### 4. Construction of Probability Distribution for Size of Individual Payments

4.1 This is step (i) from section 2. The available data on the amounts of individual payments made in the past can be used to construct the probability distribution for future payments. The main issues to be taken into account are:

- (a) The size of individual payments might be related to the stage of development. In most lines of business, the claim payments tend to be larger in later development periods than in earlier development periods. This is because small claims are settled relatively quickly. This is an issue because whereas the data will relate predominately to the earlier stages of development, future payments will fall more in later stages of development.
- (b) There might be payments in the future which are larger than any observed in the past.

The first of these (a) can be taken into account by an appropriate mixing of the empirical loss distributions constructed for each stage of development separately. This amounts to a simple re-weighting of the observed individual payment amounts. This is covered in detail in section 4.2.

The second issue (b) can be taken into account by fitting an analytic distribution with an unlimited right tail (for example, the Log-Normal or Pareto distribution). Any standard method (such as maximum likelihood estimation (MLE) or minimum distance estimation (MDE)) can be used, but a generalization is necessary to take account of the re-weighting of the individual observations mentioned above. Further details of MLE are given in section 4.3.

- 4.2.1** To construct a suitable probability distribution for the individual amounts of all future payments, taking into account point (a) of section 4.1, it is necessary to estimate what proportion of future payments will fall in each development period. To this end, the payment numbers triangle must be projected to fill in the lower right triangle and projected further to the right if necessary.

This is not covered in detail here because any suitable loss reserving method could be used — from a simple chain-ladder (link-ratio) method, to a stochastic method such as Wright (1990) — and these are documented elsewhere.

Suppose the estimated number of future payments falling in development period  $i$  is  $n_i$ . Here,  $i$  varies across development periods ( $i = 1, 2, 3 \dots$ ) and each  $n_i$  represents the total of expected future payments over all accident years. The payment amount distributions appropriate for each development period separately must be mixed together in the proportions  $p_i = n_i / N$  (where  $N$  is the total of the  $n_i$ ) in order to obtain a probability distribution appropriate for all future claim payments.

More precisely: if  $F_i(x)$  is the distribution function for the size of payments made in development period  $i$ , then the distribution function  $F(x)$  for all future payments is given by:

$$F(x) = \sum_i p_i \cdot F_i(x) \quad \text{where } p_i = n_i / N \quad (1)$$

For each  $F_i(x)$ , we can use the empirical distribution constructed from the observed (inflation adjusted) individual payments falling in each development period. That is:

$$F_i(x) = v_i(x) / v_i \quad (2)$$

where:  $v_i$  is the total observed number of payments in development period  $i$   
 $v_i(x)$  is the number of observed payments less than or equal to  $x$ .  
 $(v_i$  is an abbreviation for  $v_i(\infty)$ ).

PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY  
FROM INDIVIDUAL PAYMENTS DATA

Inserting this expression into equation (1) gives:

$$F(x) = \sum_i p_i \cdot v_i(x) / v_i$$

$$= \sum_i w_i \cdot v_i(x) / N \quad \text{where } w_i = n_i / v_i$$

Since  $n_i(x)$  is defined to be the number of observed payments in development year  $i$  which are less than or equal to  $x$ , the numerator in this expression for  $F(x)$ , is the total of the "weight"  $w_i$  for all observed payments (all development periods) which are less than or equal to  $x$ . Similarly, the denominator  $N$ , which is defined to be the total number of expected future payments, can be regarded as the total of this "weight"  $w_i$  for all observed payments:

$$N = \sum_i n_i$$

$$= \sum_i w_i \cdot v_i$$

In this way,  $F(x)$  is seen to be the empirical distribution of all observed payments (all development years combined) after applying the appropriate weight  $w_i$  to each observation.

#### 4.2.2 The simplicity of these ideas should become apparent through an example.

First, we construct the empirical distributions for each development year from the individual payment amounts given in section 3. These are the distributions  $F_i(x)$  for  $i = 1, 2, 3$ . All three empirical distributions are shown on the same set of axes in Figure 1. It is quite clear that in this example, the size of payments tends to be larger in the later development periods (this is point (a) of section 4.1).

The next step is to calculate the quantities  $v_i$  and  $n_i$  from the payment numbers triangle given in section 3. The quantities  $v_i$  are simply the column totals from the payment numbers triangle:

$$v_1 = 33, \quad v_2 = 41, \quad v_3 = 5.$$

The quantities  $n_i$  are obtained by projecting the claim numbers triangle. A simple link ratio projection, with a tail factor of 1.1, gives the following figures:

*Number of claim payments*

	Development Year			
	1	2	3	+4
Accident year 1	10	19	5	3.4
2	11	22	5.9	4.0
3	12	20.5	5.4	3.7

From which we obtain (by adding columns of the lower right triangle):

$$n_1 = 0, \quad n_2 = 20.5, \quad n_3 = 11.3, \quad n_4 = 11.1.$$

Finally, the three empirical distributions shown in Figure 1 have to be mixed according to these values for  $n_i$  in order to obtain a distribution function  $F(x)$  appropriate for all future payments.

There is a minor complication in this example because even the oldest year is not fully developed: this is why a tail factor was necessary. (The value 1.1 might have been arrived at by fitting a curve to the link ratios, or by comparison with other similar but more fully developed triangles.) This has resulted in a non-zero estimate for  $n_4$  (representing the number of future payments in the fourth and later development years) for which we have no corresponding distribution in Figure 1: we have empirical distributions only for  $i = 1, 2, 3$ .

For the purposes of this example, we will assume that the size distribution of individual payments is the same in development years 4 and later as in development year 3. Under this assumption we can treat the  $n_4$  expected payments of later development years exactly as if they were expected in development year 3:  $n_4$  is added to  $n_3$  to give an amended value:  $n_3 = 22.4$ . (Further comment on this procedure, and an outline of a more refined approach, is given in section 8.)

The total expected number  $N$  of future payments (the sum of the  $n_i$ ) is 42.9. The mixing proportions for the three empirical distributions are therefore:

$$\begin{array}{lll} p_1 = 0, & p_2 = 20.5 / 42.9 & p_3 = 22.4 / 42.9 \\ & = 0.478 & = 0.522 \end{array}$$

Mixing the two empirical distributions  $F_2(x)$  and  $F_3(x)$  in these proportions results in the distribution shown in Figure 2. As described in section 4.2.1, the mixture distribution can also be interpreted as a distribution obtained by applying a "weight" to each observed payment and pooling the observations over all development years. The weight is given by  $w_i = n_i / v_i$ , that is, for each development year, it is the ratio of number of payments expected in the future, to number observed in the past. For the present example, this gives:

$$\begin{array}{lll} w_1 = 0, & w_2 = 20.5 / 41 & w_3 = 22.4 / 5 \\ & = 0.478 & = 4.48 \end{array}$$

These weights have a simple intuitive interpretation: the 41 observations for development year 2 represent an expected number of 20.5 future claim payments in development year 2, so each receives a weight of  $20.5 / 41$ . The 22.4 expected claims in later development years are represented by only 5 in the data, so each of these 5 receives a weight of  $22.4/5$ .

**4.3.1** Having constructed an appropriate empirical loss distribution as described in section 4.2, the next step is to smooth it and project a right tail to allow for the

PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY  
FROM INDIVIDUAL PAYMENTS DATA

possibility of future payments larger than any observed in the past. We will use maximum likelihood estimation (MLE), adjusted to allow for the weights  $w_i$ . The empirical distribution can be regarded as a number of data-pairs  $(x_j, w_j)$ , where  $x_j$  is an observed payment amount, and  $w_j$  is the corresponding weight. (The weight  $w_j$  is the same for all observations  $x_j$  falling in the same development year  $i$ , because  $w_j = n_i / v_i$ , where  $i$  is the development year of observation  $j$ .)

We will use  $f(x|\beta)$  to denote the probability density function of the analytic distribution to be fitted to the data.  $x$  represents the amount of an individual payment as before, and  $\beta$  represents the parameters of the analytic family (in general  $\beta$  will be a vector). In the absence of weights, MLE means setting the parameters  $b$  to those values which maximise the log-likelihood  $L$ , which is given by:

$$L = \sum_j \ln(f(x_j|\beta)) \quad (3)$$

If our data included repeated values  $x_j$ , the same term  $\ln(f(x_j|\beta))$  would appear more than once in the summation. Therefore, repeated  $x_j$  values could be removed by keeping a count  $r_j$  of the number of repeats of each distinct  $x_j$  value ( $r_j$  would be 1 for each unique  $x_j$  value). The log-likelihood would then be given by:

$$L = \sum_j r_j \cdot \ln(f(x_j|\beta)) \quad (4)$$

Clearly, the higher the  $r_j$  corresponding to an  $x_j$ , the greater the influence of that  $x_j$  value in the fit obtained by maximising  $L$ : for this reason, the  $r_j$  can be regarded intuitively as "weights". This explains how integer weights might arise. For our purposes, the weights  $w_j$  could have non-integer values, but nevertheless, the weight should determine the influence of each observation  $x_j$  in determining the analytic distribution and this is achieved by using them as multipliers of each term in the objective function  $L$  to be maximised. For this reason, the appropriate generalisation to basic MLE (equation (3) above) is:

$$L = \sum_j w_j \cdot \ln(f(x_j|\beta)) \quad (5)$$

Given a specific formula for  $f(x|\beta)$  (for example, the Pareto or Log-Normal formula), this optimisation problem can be solved using a numeric method such as the Newton-Raphson algorithm in exactly the same way as in basic distribution fitting by MLE (equation (3)). The only difference is in the interpretation of standard errors of the resulting parameter estimates: further discussion on this aspect is given in section 8.

**4.3.2** The table below gives the data-pairs  $(x_j, w_j)$  for our example data-set. The first column is the payment amount  $x_j$  (this is taken directly from section 3 for all those payments made in development years 2 and 3). The second column gives the weight  $w_j$ . This is 0.478 or 4.48, for payments made in development years 2 and 3 respectively (see section 4.2.2):

1268	0.478	2225	0.478	2323	0.478	3135	0.478
3639	0.478	3815	0.478	3953	0.478	4187	0.478
4792	0.478	4960	0.478	5212	0.478	5253	0.478
5395	0.478	5953	0.478	6094	0.478	6210	0.478
6251	0.478	6319	0.478	6401	0.478	6615	0.478
6801	0.478	6905	0.478	7048	0.478	7062	0.478
7418	0.478	7589	0.478	7647	0.478	7886	0.478
8573	0.478	8661	0.478	9244	0.478	9515	0.478
9553	4.480	10043	0.478	10160	0.478	11826	0.478
11905	0.478	12968	0.478	13265	0.478	13609	4.480
16659	0.478	16756	0.478	18455	4.480	19875	0.478
22278	4.480	34091	4.480				

Several different distribution families  $f(x|\beta)$  have been fitted to these data by generalized MLE (equation (5) of section 4.3.1). The following two-parameter distribution families all provide reasonably good fits:

Gamma:

$$f(x|\beta) = \frac{1}{\beta_1 \times \Gamma(\beta_2)} \times \left(\frac{x}{\beta_1}\right)^{\beta_2-1} \times \exp\left(-\frac{x}{\beta_1}\right)$$

Weibull:

$$f(x|\beta) = \frac{\beta_2}{\beta_1} \times \left(\frac{x}{\beta_1}\right)^{\beta_2-1} \times \exp\left(-\left(\frac{x}{\beta_1}\right)^{\beta_2}\right)$$

Log-Normal:

$$f(x|\beta) = \frac{1}{x \times \sqrt{2 \times \pi \times \beta_2}} \times \exp\left(-\frac{1}{2 \times \beta_2} \times \log^2\left(\frac{x}{\beta_1}\right)\right)$$



PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY  
FROM INDIVIDUAL PAYMENTS DATA

The maximized log-likelihood values, and corresponding parameter estimates, are given in the following table:

<i>Family</i>	<i>Log-likelihood</i>	$\beta_1$	$\beta_2$
Gamma	-444.552	5,742.2	2.41490
Weibull	-445.140	15,570.2	1.63220
Log-Normal	-445.184	11,115.9	0.48887

On the basis of the log-likelihood, the Gamma distribution appears to provide the best fit (largest value of L). However, when assessing the fit of analytic distributions to data, it is advisable also to look at graphs. All three of these fitted distributions are shown in Figure 3, in each case, on the same set of axes as the empirical distribution (from Figure 2). These graphs suggest that for the small and medium sized claim payments (up to about £14,000) the Log-Normal distribution perhaps provides a better fit than the Gamma.

Another method for assessing the fit of analytic distributions graphically, is to look at the mean residual life function. This method is recommended by Hogg & Klugman (1984). The mean residual life function is defined by:

$$\text{mrl}(x) = \frac{E(\max(0, X - x))}{1 - F(x)}$$

where  $X$  is the random variable, and  $x$  is any fixed value of  $X$ . That is,  $\text{mrl}(x)$  is defined to be the mean amount in excess of  $x$ , of payments which do exceed  $x$ .

Figure 4 shows this function plotted for each of the three fitted analytic distributions, in each case on the same set of axes as the empirical mean residual life. For this purpose, the empirical mean residual life has been calculated directly from the data-points  $(x_j, w_j)$  using the formula:

$$\text{mrl}(x) = \frac{\sum w_j \times (x_j - x)}{\sum w_j}$$

where summation, in both numerator and denominator, is over only those observed payment amounts  $x_j$  which exceed  $x$ .

Assessing the quality of fit in this way supports the indications of the log-likelihood values: the Gamma distribution seems to provide the closest fit to the data. We will continue this example in later sections using only the Gamma distribution for individual future payments. In practice, it would be advisable to carry out the remaining calculations for at least two different fitted distributions and to compare the results.

## 5. Construction of Probability Distribution for the Number of Future Payments

- 5.1 This is step (ii) from section 2. The simplest approach is to model and project the payment numbers triangle. This can be carried out using almost any aggregate reserving method. If a stochastic method is used, this will provide a standard deviation as well as an expectation for the number of future payments. If a non-stochastic method is used, a simple "rule of thumb" for obtaining a suitable value for the standard deviation is to take the variance to be two times the expectation. More refined methods might take account of the reported claim numbers triangle and other relevant information.

Having obtained, one way or another, an expected value and standard deviation for the number of future payments, the complete distribution for the number of future payments can be taken as Negative Binomial with parameters obtained by equating the first two moments. There are two reasons for using the Negative Binomial distribution, the first theoretical and the second practical:

- (i) The number of future payments is a counting process. The number of payments arising from each small part of the cohort separately (for example, each individual policy), could reasonably be modelled as Poisson. If all these component parts were stochastically independent, then the total number of future payments would also be Poisson (as the sum of independent Poissons). However, the component parts will not be stochastically independent for a variety of reasons (for example, the possibility of a legal decision which sets a precedent). The total number is in fact an aggregation of correlated Poisson variables. Such an aggregation is usually well approximated as Negative Binomial. It is exactly Negative Binomial if the Poisson parameters of the component parts are related through a common random factor with a Gamma distribution. The Gamma distribution is sufficiently flexible for this theoretical result to be useful in practice. Parameter uncertainty is another element which can reasonably be taken into account by using a Negative Binomial distribution instead of a Poisson. For further detail on these issues, see for example Heckman & Meyers (1983).
- (ii) The calculation of the distribution of aggregate outstanding liabilities (step (iii) from section 2) is easiest if the count distribution is Negative Binomial. Both Panjer's algorithm and Heckman/Meyers algorithm require the count distribution to be from this family. (In fact, both algorithms work also with the Poisson and Binomial distributions, but these distributions are not appropriate for future claim numbers because they do not have variance greater than mean, which is inevitable for the reasons mentioned under (i) above.)

# PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY FROM INDIVIDUAL PAYMENTS DATA

The basic Negative Binomial distribution is the distribution of the number of failures before success number  $m$  in independent binomial trials each with probability  $p$  of success. It is given by:

$$\text{Prob}(N = r | m, p) = \frac{(r + m - 1)!}{r! \times (m - 1)!} \times p^m \times (1 - p)^r$$

However, this can be generalised to give a distribution for any  $m > -1$  (not restricted to integer values) by noting that:

$$\text{Prob}(N = r) = (1 - p) \times \left( 1 + \frac{m - 1}{r} \right) \times \text{Prob}(N = r - 1)$$

and calculating the distribution recursively. (This is a case of Panjer's recursive formula.)

Given the first two moments,  $E(N)$  and  $\text{Var}(N)$ , the parameters of the Negative Binomial are given by:

$$p = E(N) / \text{Var}(N) \qquad m = (p / (1 - p)) \cdot E(N). \qquad (6)$$

- 5.2 In this section, we continue the example of sections 3 and 4.2.2. In section 4.2.2 we projected the payment numbers triangle from section 3 using the chain ladder method with a tail-factor of 1.1. These gave an expectation of 42.9 future payments. Applying our rule of thumb gives a prediction variance of 85.8, hence a standard deviation of 9.26. Putting these values of the first two moments in to equations (6) gives the following parameters for the Negative Binomial distribution:

$$p = 0.5 \qquad m = 42.9.$$

This Negative Binomial distribution has a skewness coefficient of 0.3237 and implies 80% confidence that the number of future payments will be in the range 32 to 54 inclusive.

## 6. Calculation of Compound Distribution

As stated in section 2, this stage is not covered in detail in this article as several good methods are documented elsewhere. We briefly describe the use of Panjer's recursive method for the example of earlier sections.

To apply Panjer's method, the individual payment distribution must be discrete rather than continuous. In section 4.3.2 a Gamma distribution was found for the distribution of future payment amounts. Before applying Panjer's algorithm, this must be approximated as a discrete distribution. The approximation used for this example is shown in Figure 5. This has a step-width of £500, and each step height in Figure 5 is the probability that an individual payment will be equal to the

corresponding multiple of £500. This discretization was achieved by finding the average value of the original Gamma cumulative distribution function within each band of width £500. This was continued up to £100,000, giving a total of 201 discrete probabilities. The Gamma distribution indicated a chance of only  $1.345 \times 10^{-6}$  that an individual payment would exceed £100,000.

Having discretized the payment amount distribution, it can be combined with the Negative Binomial distribution for the number of outstanding payments (derived in section 5.2) in order to calculate the distribution of the total outstanding liability (refer to Panjer, 1981 for details). Figure 6 shows both the cumulative distribution function and the probability density function of the result.

The mean of this distribution is £594,880: this is simply the product of the mean £13,866.80 of the Gamma distribution, and the expected number 42.9 of outstanding payments. The standard deviation is £141,102. It is clear from the density function in Figure 6 that the outstanding liability is positively skewed: this is due partly to the skewness of the Gamma distribution for individual payments, and partly to the skewness of the Negative Binomial distribution for the number of payments. The coefficient of skewness is 0.359.

## 7. Setting the Reserve — Safety Loads

As already pointed out in section 1, many stochastic reserving methods yield only an expectation and a standard deviation for the future liability. To make an allowance for possible adverse future run-off, the reserve might be set by adding some multiple of the standard deviation (for example, one standard deviation) to the expected value. In the present example, adding one times the standard deviation gives a reserve of £735,982. But what is the chance that this reserve will be inadequate, and by how much is it likely to be exceeded if it does prove inadequate? To answer these questions, the more complete information illustrated in Figure 6 is necessary. From Figure 6, the chance that the outstanding liability will exceed £735,982 is 15.5%. The mean excess should this occur is the mean residual life of the distribution evaluated at £735,982: this is £86,683.

Given the more complete information represented by Figure 6, there are many other options for setting the reserve. We could use the criterion that there must be a certain probability (for example 90%) that the reserve will prove to be adequate. This means setting the reserve to a percentile of the aggregate loss distribution: from Figure 6, the reserve with a 90% chance of adequacy is £780,000. However, this criterion does not allow for the potential magnitude of the inadequacy in the (10%) worst cases.

A safety load criterion which does make such an allowance, and which has several other desirable theoretical properties, is the proportional hazards criterion (Wang, 1995). This was suggested by Wang in the context of reinsurance pricing, but is equally valid for safety loads in reserves. The safety loaded premium is calculated as the mean of a distribution obtained by transforming the distribution of the total outstanding liability. The transformation suggested by Wang, the "proportional

hazards" transform, is the raising of the survivor function  $1 - F(t)$  to some power  $1/\delta$ , where  $\delta$  is an index greater than 1 reflecting the degree of caution required. This is illustrated by Figure 7, which shows both the outstanding liability distribution (as in Figure 6), and its proportional hazards transform with  $\delta = 1.5$ .

The mean of this transformed distribution is £653,677: this would be the reserve under this criterion. Note that, because the mean of any probability distribution is equal to the area above the cumulative distribution function, the safety load of £58,797 (the difference in the two means: £653,677 – £594,880), is represented by the area between the two curves in Figure 7. The higher the index  $\delta$ , the larger this area would become. Using an index  $\delta = 2$  gives a reserve of £702,821 (safety load of £107,941) and using  $\delta = 3$  gives £784,786 (safety load of £189,906), which is just above the 90th percentile found earlier.

## 8. Conclusion

The aim of this paper is to present a general approach to claims reserving rather than a single method. The essence of the approach is to make full use of detailed claim payment information in order to calculate a complete probability distribution for the outstanding liability. This approach has been made possible by parallel advances in computer technology and actuarial theory. It is hoped that the ideas presented in this article will provide a framework within which these advances can be turned to advantage in the reserving context.

The example presented in this paper illustrates the approach but is not intended as a model suitable for general application. Some of the more obvious refinements and alternatives which may be necessary in practice are mentioned briefly in the remainder of this section.

In the absence of any observed payments for the fourth and later development periods in the example data-set, we assumed that their distribution would be the same as for the third development period. An alternative would be to project trends observed in the distribution across development periods. This is most easily achieved by fitting an analytic distribution before mixing the development years rather than after.

In the example, we could have tried fitting a Gamma distribution to each of the three empirical distributions shown in Figure 1 separately. This would quite likely indicate a trend increase in the  $\beta_1$  parameter across development periods 1, 2, 3, and possibly a trend in  $\beta_2$  as well. These trends could be projected to arrive at a suitable Gamma distribution for development periods four and later. Each of the Gamma distributions could then be discretized before mixing them to obtain a distribution suitable for all future payments.

The approach can also be refined to take account of uncertain future claims inflation and investment returns, and uncertainty in the parameters of the distributions for the number and amount of individual losses.

The rule of thumb used in the example for determining the variance of the Negative Binomial distribution for claim numbers is very crude. In practice, it is worth taking at least as much trouble modelling the payment number distribution as the payment size distribution, because when the expected number of payments is large, the payment size distribution becomes relatively less influential. It is recommended that a stochastic method should be used to project the claim numbers triangle, and that the variance of the Negative Binomial be set to reflect both the future process variance and the parameter uncertainty indicated by the stochastic method. The rule of thumb presented in this paper arises from the observation that when claim number triangles are modelled thoroughly in this way, the resulting prediction variance does quite often turn out to be around two times the best estimate of the number of future payments.

Finally, a word of caution about the interpretation of the parameter standard errors resulting from distribution fitting to weighted observations using the generalised version of MLE as described in section 4.3.

If the Newton-Raphson method is used for MLE, a spin-off is that the inverse of the final Hessian matrix is approximately the variance-covariance matrix of the parameter estimates. This reflects both the volume of data, and the quality of the fit to that data. However, when the observations are re-weighted as illustrated in the example, the total "number of observations" becomes the sum of the weights, and the scale of the variance-covariance matrix will depend on this quantity.

If the variance-covariance matrix is required, it is probably safest to rescale all weights so that their total remains equal to the true number of observations. In section 4.3.2, there are 46 observations with a total weight of 42.9. In this case, the weights should be scaled up by the factor 1.07226 before fitting, giving revised weights of 0.513 and 4.804, for development periods 2 and 3 respectively.

## References

Heckman P E & Meyers G G (1983): The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions, PCAS.

Hogg R V & Klugman S A (1984): Loss Distributions, Wiley, New York.

Panjer H H (1981): Recursive Evaluation of Compound Distributions, ASTIN 12, pp 22–26.

Wang S S (1995): Insurance Pricing and Increased Limits Ratemaking by Proportional Hazards Transforms, Insurance: Mathematics and Economics.

Wright T S (1990): A Stochastic Method for Claims Reserving in General Insurance, JIA Vol. 117, part III.

PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY  
FROM INDIVIDUAL PAYMENTS DATA

Figure 1: Empirical Distribution for Each Development Year

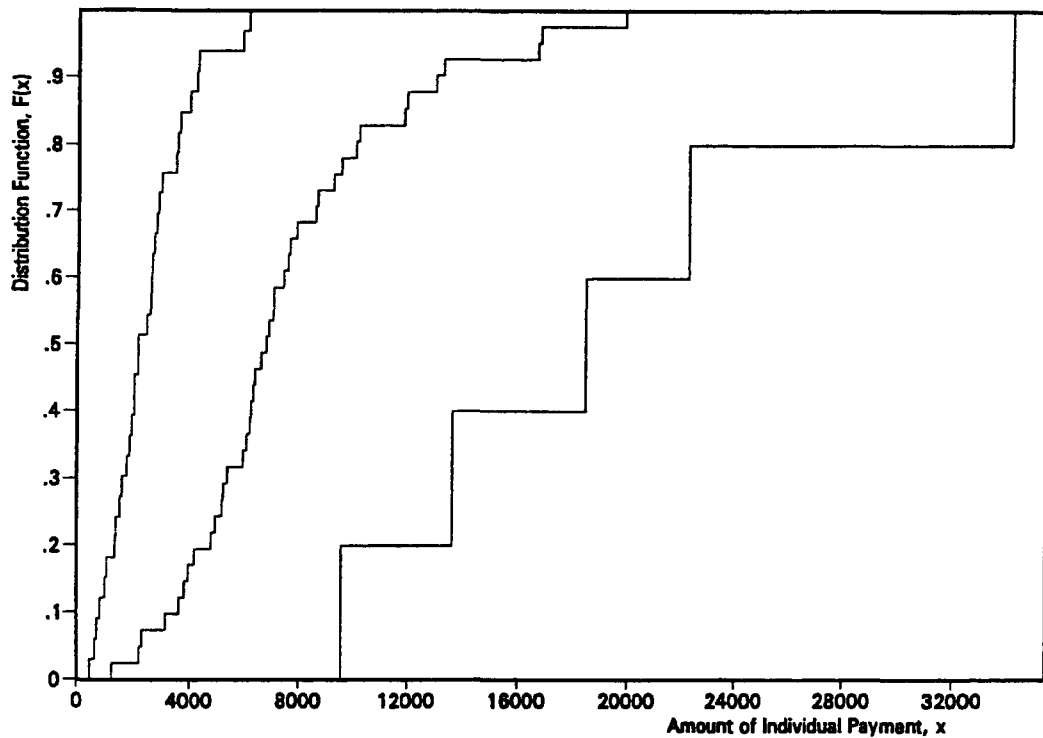


Figure 2: Mixture of Development Years' 2 and 3 Distributions  
(appropriate for all future claim payments)

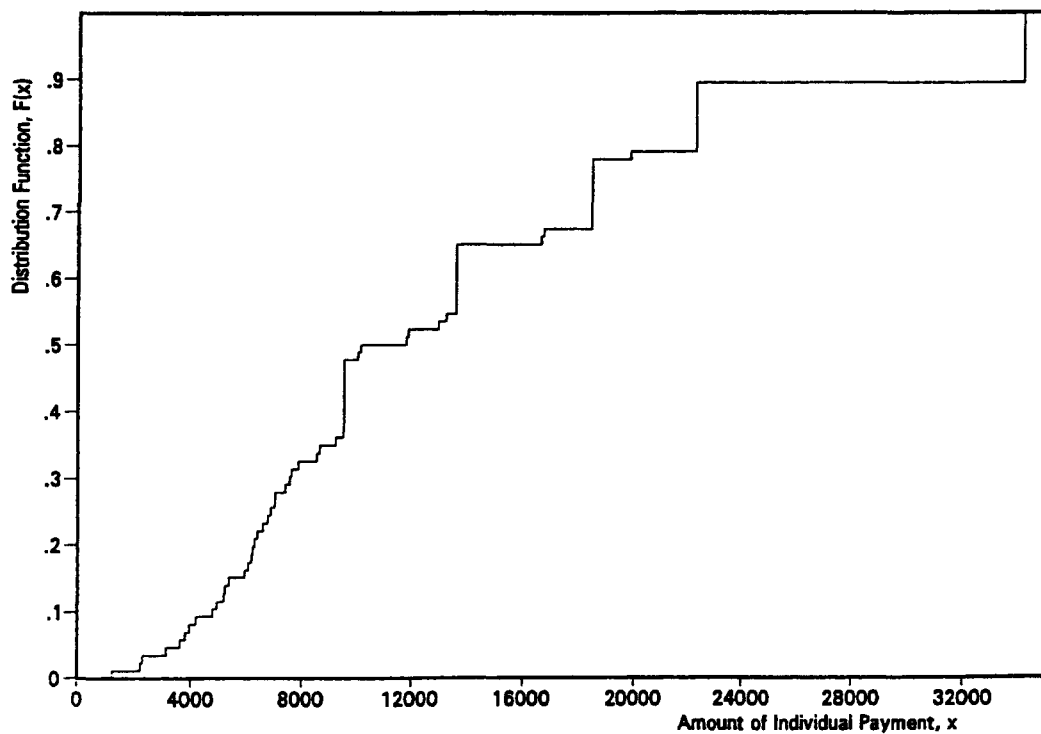


Figure 3a: Fitted Gamma Distribution for Future Payments

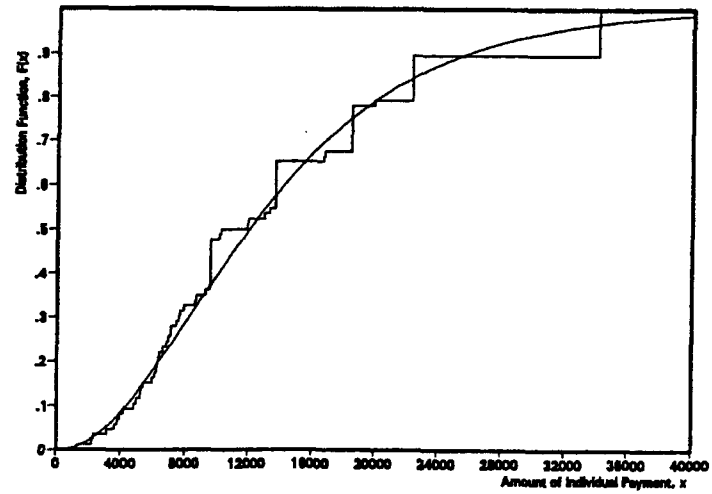


Figure 3b: Fitted Weibull Distribution for Future Payments

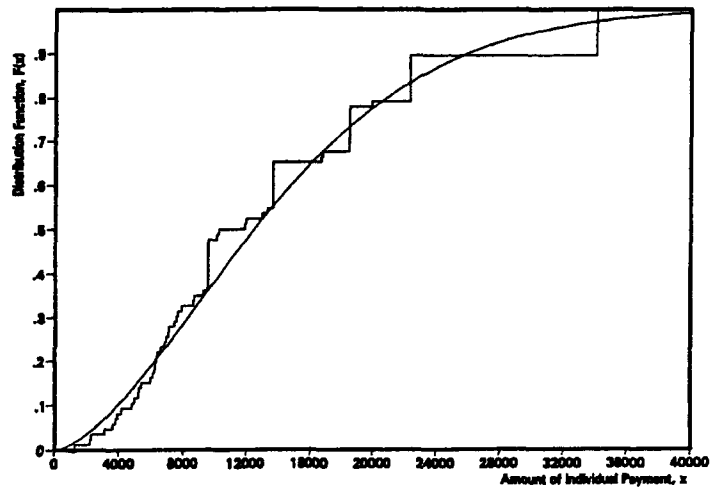
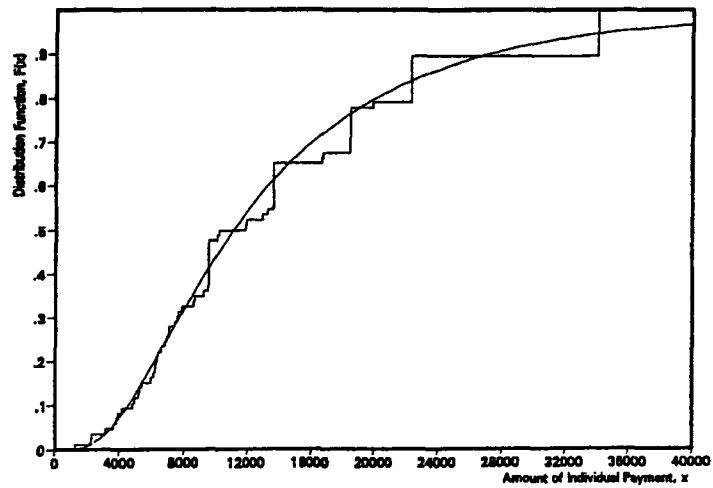


Figure 3c: Fitted Log-Normal Distribution for Future Payments





PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY  
FROM INDIVIDUAL PAYMENTS DATA

Figure 4a: Fitted Gamma Distribution for Future Payments

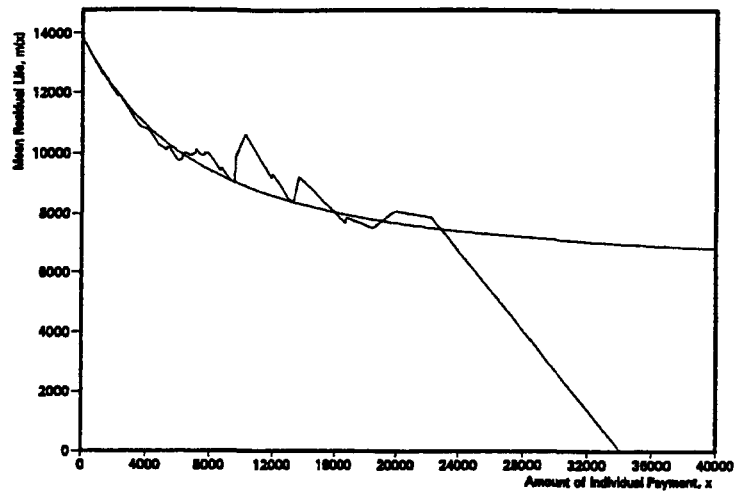


Figure 4b: Fitted Weibull Distribution for Future Payments

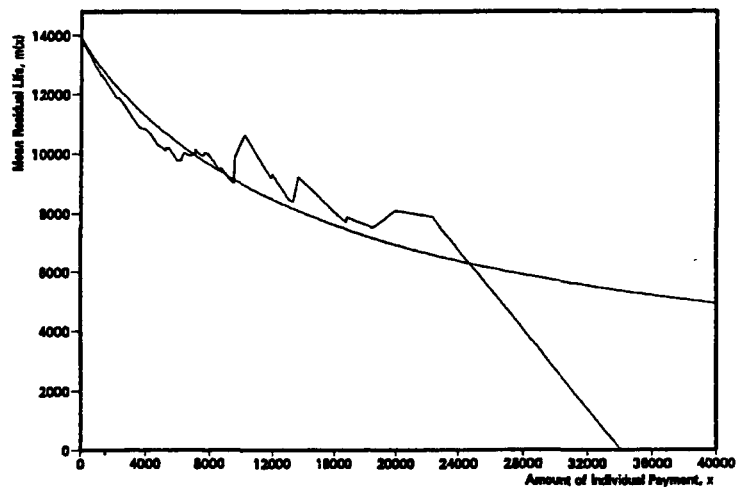
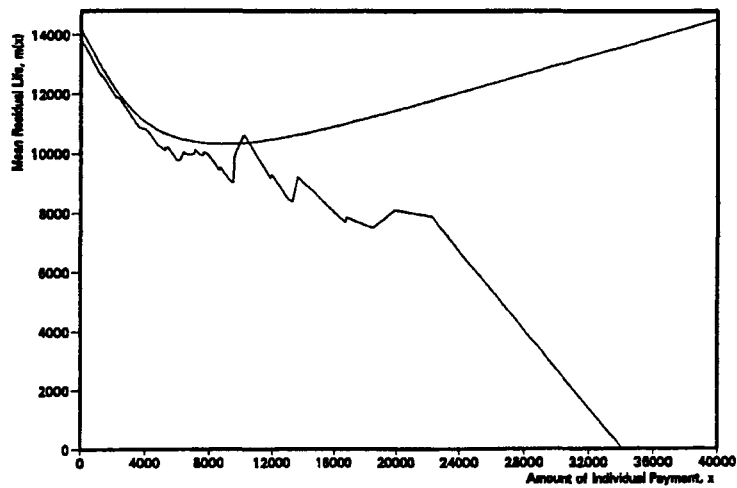
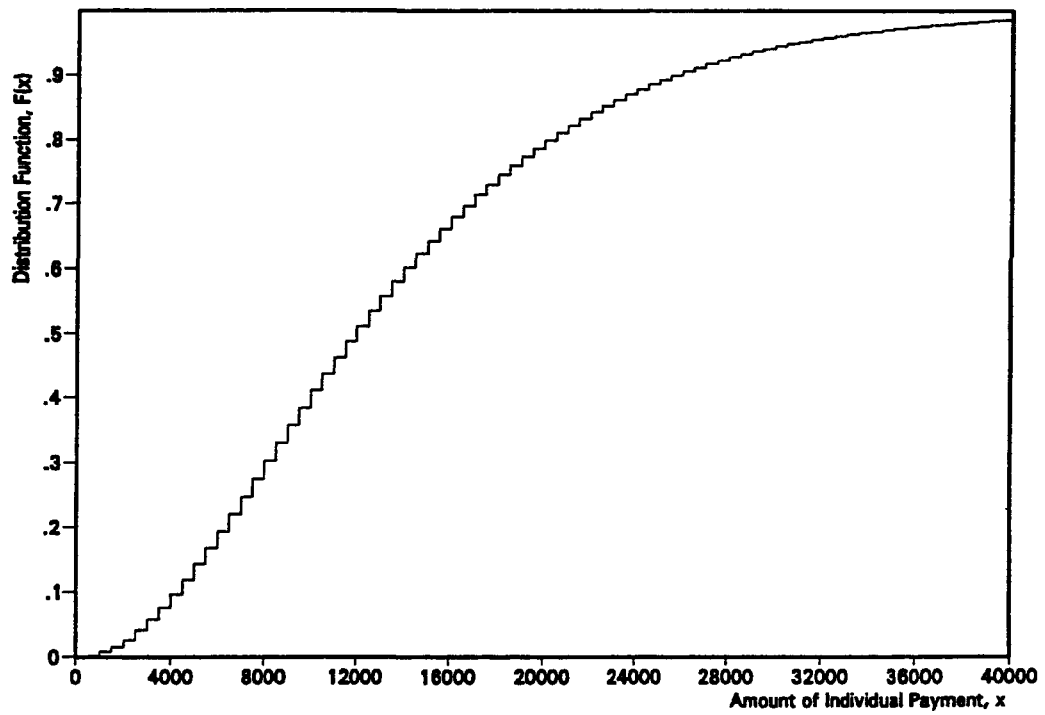


Figure 4c: Fitted Log-Normal Distribution for Future Payments

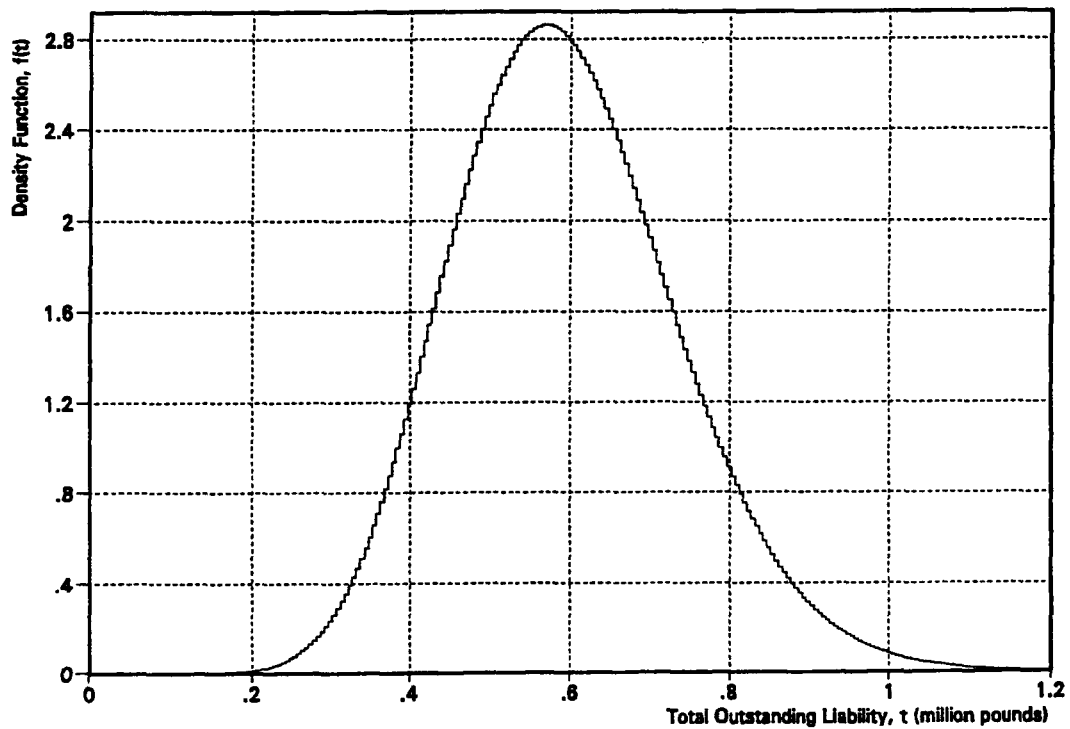
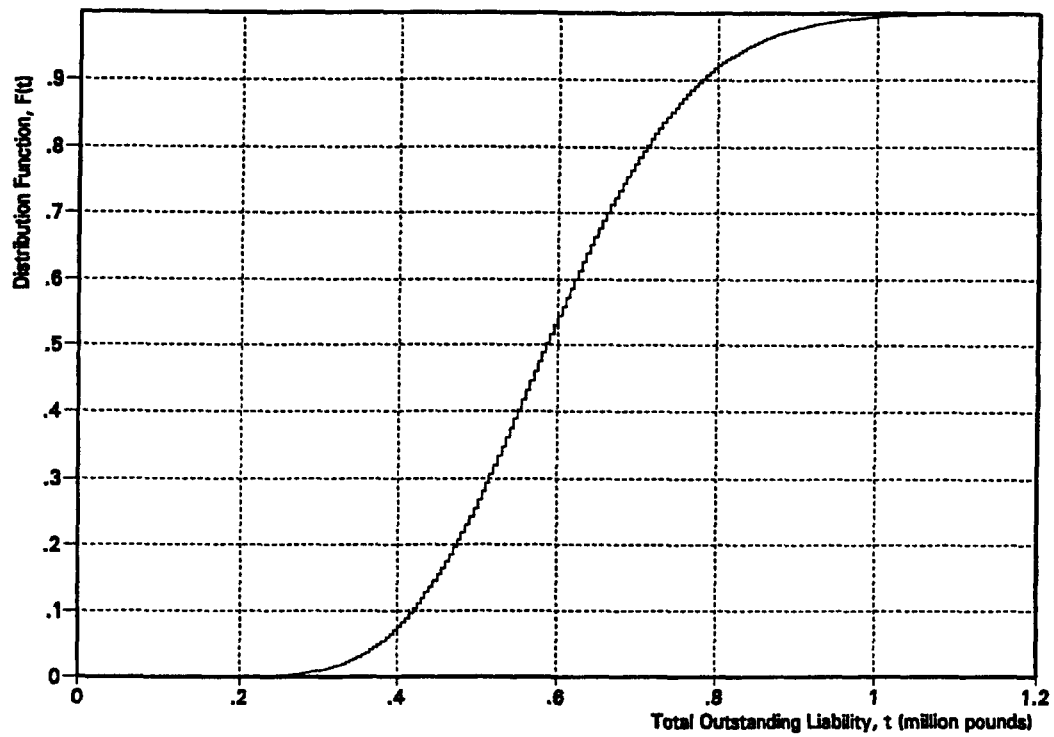


**Figure 5: Discretized Gamma Distribution for Future Payments**

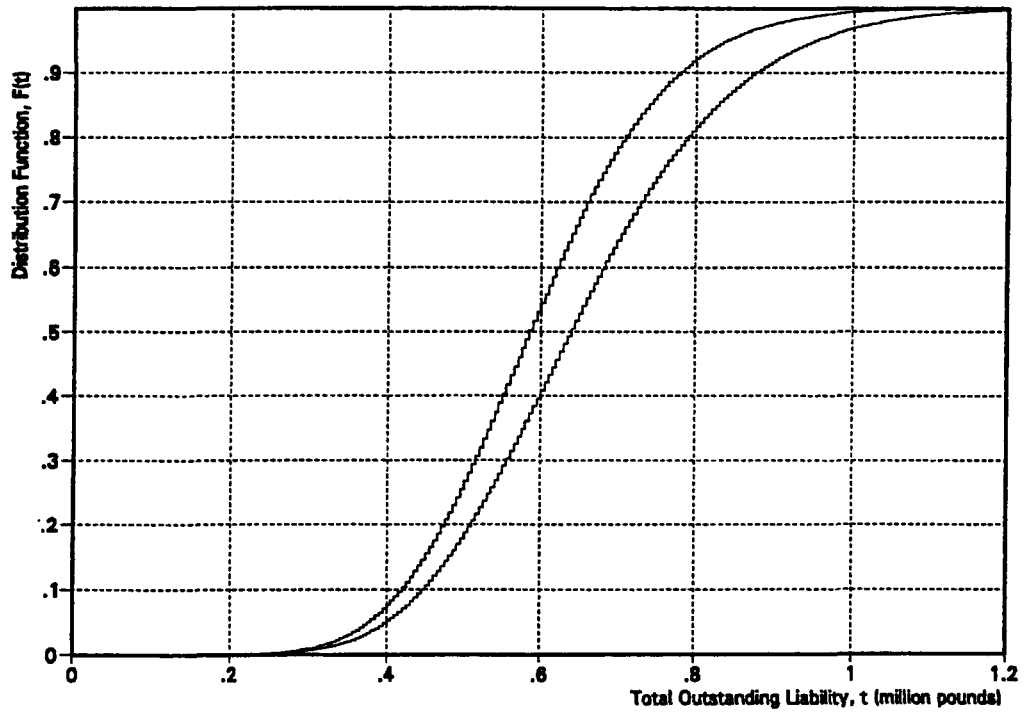


PROBABILITY DISTRIBUTION OF OUTSTANDING LIABILITY  
FROM INDIVIDUAL PAYMENTS DATA

Figure 6: Probability Distribution of Outstanding Liability



**Figure 7: Proportional Hazards Transform with Index = 1.5**



◇