THE UNIVERSITY OF CALGARY

On the Computation of Total Claims Distributions

by

Abu Bashar Md. Shafiqur Rahman

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

August, 2002

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "On the Computation of Total Claims Distributions" submitted by Abu Bashar Md. Shafiqur Rahman in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

R. S. Ambaga

Supervisor, Dr. R. Ambagaspitiya
Department of Mathematics and Statistics

M. D. Burke

Dr. M. D. Burke
Department of Mathematics and Statistics

Dr. A. Sesay
Department of Electrical and Computer Engineering

August 29 2002

Date

Abstract

In this thesis we implement a recursive algorithm for computing total claims distribution with generalized poisson claim counts in Visual Basic For Application. The program that we implement can be used as an add-in to Excel. The intended audience is practicing actuaries in the property and casualty area. They use Excel extensively in pricing insurance policies.

In the introduction we discuss total claims distribution. We also present an extensive review of claim counts models and some results on claim sizes modelling. We briefly review techniques available to compute total claims distributions outlining advantageous and disadvantageous.

Acknowledgements

I wish to express my deep and sincere gratitude to my supervisor Dr. Rohana S Ambagaspitiya, for his supervision, encouragement and guidance during the preparation of this thesis. I also gratefully acknowledge his financial support during my research. I would like to thank to the examination committee for their valuable suggestions.

Also I acknowledge the financial support of the department of Mathematics and Statistics, University of Calgary. I am very grateful to all teaching and support staff of this department for their help and cooperation during my course of study.

Finally, I express appreciation to my parents, my wife Tasneem Kawsar, daughter Izwani shafiq and other family member for their encouragement and inspiration.

Table of Contents

List of Tables

Chapter 1

Introduction

In the collective risk model the basic concept is that of a random process that generates claims for the portfolio of policies. This process is characterized in terms of the portfolio as a whole rather than in terms of the individuals policies comprising the portfolio.The mathematical formulation is as follows: Let N denote the number of claims produced by a portfolio of policies in a given time period. Let X_1 denote the amount of the first claim, X_2 the amount of the second claim and so on. Then

$$S = X_1 + X_2 + ... + X_N \tag{1.1}$$

represents the aggregate claims generated by the portfolio for the period under study. The number of claims, N, is a random variable and is associated with the frequency of claim. In addition, the individual claim amounts $X_1, X_2, ...$ are also random variables and are said to measure the severity of claims.

In order to make the model tractable, two fundamental assumptions are made in the actuarial literature. These are

1. $X_1, X_2, ...,$ are identically distributed random variables.

2. The random variables $N, X_1, X_2, ...$ are mutually independent.

The expression (1.1) will be called a random sum, and assumptions (1) and (2) will always be made concerning its components.

A first step in exploring this alternative model will be the study of the distribution of S in terms of the distribution of N and of the common distribution of the X_i's.

A second step, is the discussion of choices for the distribution of N and the common distribution of the X_i's. For N a poisson or a negative binomial distribution is often selected. For the claim amount distribution, a normal, gamma or other, perhaps empirical, distribution may be used. When a Poisson distribution is chosen for N, the distribution of S is called a compound Poisson distribution; when a negative binomial distribution is selected for N, the distribution of S is called a compound negative binomial distribution.

The distribution of aggregate claims in a fixed time period can be obtained from the distribution of the number of claims and the distribution of individual claim amounts.

Let $P(x)$ denote the common distribution function (df) of the independent and identically distributed X_i's. Let X be a random variable with this d.f.; let

$$p_k = E[X^k] \tag{1.2}$$

denote the kth moment about the origin and

$$M_X(t) = E[e^{tX}] \tag{1.3}$$

the moment generating function (mgf) of X. In addition, let

$$M_N(t) = E[e^{tN}] \tag{1.4}$$

denote the mgf of the distribution of number of claims, and let

$$M_S(t) = E[e^{tS}] \tag{1.5}$$

denote the mgf of aggregate claims. The df of aggregate claims will be denoted by $F_S(x)$.

There are some general formulas relating the moments of random variables by conditional expection. For mean and variance these are

$$E[W] = E[E[W|V]] \tag{1.6}$$

$$Var[W] = Var[E[W|V]] + E[Var[W|V]] \tag{1.7}$$

Using (1.6) and (1.7), in conjunction with assumption(1) and (2) we obtain

$$E(S) = E[E[S|N]] = p_1 E[N] \tag{1.8}$$

and

$$\begin{aligned} Var[S] &= E[Var[S|N]] + Var[E[S|N]] \\ &= E[NVar[X]] + Var[p_1 N] \\ &= E[N]Var[X] + p_1^2 Var[N] \end{aligned} \tag{1.9}$$

where $Var[X] = p_2 - p_1^2$

The result stated in (1.8), that the expected value of aggregate claims is the product of the expected individual claim amount and the expected number of claims, is not surprising. Expression (1.9) for the variance of aggregate claims also has a natural interpretation. The variance of aggregate claims is the sum of two components where the first is attributed to the variability of individual claim amounts and the second to the variability of the number of claims.

In a similar fashion we derive an expression for the mgf of S :

$$M_S(t) = E[e^{tS}] = E[E[e^{tS}|N]]$$

$$= E[M_X(t)^N] = E[e^{N \log M_X(t)}]$$
$$= M_N(\log M_X(t)) \tag{1.10}$$

In Chapter 2 we present an extensive discussion of models that have been proposed as claims counts distributions. In Chapter 3 we discuss modelling claim sizes. In Chapter 4 we review the methods available in the literature to compute claims distributions. Also in Chapter 4 we discuss some implementation issues in visual basic for applications (VBA) of a recursive algorithm. Finally, we present a selected sample of outputs and the VBA codes.

Chapter 2

Models for Claim Count

2.1 Introduction

It is obvious that claim counts have to be modeled by discrete random variables with nonnegative support. In the Statistical literature we may find a large number of discrete distributions. Johnson, Kotz and Kemp (1992) is a book length account of univariate discrete distributions. However, actuaries used only a subset of these distributions to model claim counts; discrete distributions that yield easily computable total claims distributions have been suggested as alternatives to the Poisson distribution.

In this chapter we represent a comprehensive survey of the discrete distributions that have been used to model claim counts. We divide them into four broad classes. The (a, b) class distributions, compound distributions, mixture distributions and the class$(a + b, b)$ distributions.

2.2 The (a, b) Class of Distribution

Consider the class of counting distributions (with support on the non-negative integers) for which the recurrence relation

$$\frac{p_n}{p_{n-1}} = a + \frac{b}{n}, \qquad n = 1, 2, 3..... \tag{2.1}$$

holds. In the actuarial literature, this class is identified as class (a,b) distributions by Sundt and Jewell (1981). This recursion describe the relative size of successive probabilities in the counting distribution. The probability at zero, p_0, can be obtained from the recursive formula since the probabilities must add up to 1. This provides a boundary condition. The (a,b) class of distributions is a two-parameter class, the two parameters being a and b. By substituting in the probability function for each of the Poisson, binomial, and negative binomial distributions on the left hand side of the recursion, it can be seen that each of these three distributions satisfies the recursion and that values of a and b are as given in the table. In addition the table gives the value of p_0, the starting value for the recursion.

Table 2.1: Recurrence relation

Distribution	a	b	p_0
Poisson	0	λ	$\exp(-\lambda)$
Binomial	$-\frac{q}{1-q}$	$(n_1+1)\frac{q}{1-q}$	$(1-q)_1^n$
Negative Binomial	$\frac{\beta}{1+\beta}$	$(r-1)\frac{\beta}{1+\beta}$	$(1+\beta)^{-r}$

Any distribution satisfying the following recurrence relation

$$\frac{p_n}{p_{n-1}} = a + \frac{b}{n}, \qquad n = m, m+1, ... \tag{2.2}$$

is called the (a, b, m) class distribution. When $m = 2$, the resulting family is called zero modified class (a, b) distributions.

2.3 Compound Distribution

The history of compound distributions is very large. Neyman(1939) constructed a statistical models of the distribution of larvae in a unit area of a field (in a unit of habitat) by assuming that the variation in the number of clusters of eggs per unit area (per unit of habitat) could be represented by a poisson distribution with parameter λ, while the numbers of larvae developing from the cluster of eggs are assumed to have independent poisson distributions all with same parameter ϕ.

Consider the initial (zero) and the first generations of a branching process. Let the probability generating function (pgf) for the size M of the initial (parent) generation be $P_1(z)$, and suppose that each individual i of this initial generation independently give rise to a random number Y_i of first generation individuals, where $Y_1, Y_2, ...$ have a common distribution, that of Y with pgf $P_2(z)$. The random variable for the total number of the first generation individuals is then

$N = Y_1 + Y_2 + ... + Y_M$,

where N and $Y_i, i = 1, 2, ..., M$ are all random variables. The probability generating function (pgf) of the distribution of S_N is

$$E[z^N] = E_M\{E[z^N|M]\} \tag{2.3}$$

since $E[z^N|M]$ means expected value of the product of N independent random variables, we have

$$E[z^N|M] = [P_2(z)]^M$$

$$\text{Therefore, } E[z^N] = E_M\{[P_2(z)]^M\} = P_1(P_2(z))$$

$$P_N(z) = P_1(P_2(z)) \tag{2.4}$$

Some other researchers have described $P_1(z)$ as primary distribution, $P_2(z)$ as secondary distributions and distributions with pgf's of the form $P_1(P_2(z))$ "compound". A compound distribution is a combination of two independent distributions. This process was called "generalized" by Feller(1943). Some authors (for example Douglas, 1971,1980) have chosen to use term "stopped-sum" instead for this type of distribution, because the principle model for the process can be interpreted as the summation of observations from the distribution F_2, where the number of observations to be summed is determined by an observation that is stopped by the value of the F_1 observation. Another term chosen by some other authors for this type of distribution is "contagious" distribution.

Let us look at an example of employing compound claim frequency distribution in a study of claims on automobile insurance. The primary distribution describing the number of accidents in a fixed time period is assumed to be poisson and the secondary distribution describing the number of claims per accident is assumed to be logarithmic. i.e.

$$\begin{aligned}
P_1(z) &= \exp\left(\lambda(z-1)\right) \\
P_2(z) &= \frac{\log[1-\beta(z-1)] - \log(1+\beta)}{\log(1+\beta)} \\
&\text{where} \qquad \beta = \frac{q}{p}
\end{aligned}$$

$$\begin{aligned}
P_N(z) &= P_1(P_2(z)) \\
&= \exp\left[\lambda\left[-\frac{\log[1-\beta(z-1)]}{\log(1+\beta)} + 1 - 1\right]\right] \qquad (2.5) \\
&= \exp\left(-\log\left[(1-\beta(z-1))^{\frac{\lambda}{\log(1+\beta)}}\right]\right)
\end{aligned}$$

$$= \frac{1}{[1-\beta(z-1)]^{\frac{\lambda}{\log(1+\beta)}}}$$
$$= \frac{1}{[1-\beta(z-1)]^{r}}$$

Where, $\quad r = \dfrac{\lambda}{\log(1+\beta)} > 0$

The resulting distribution of the number of claims in the fixed time period is negative binomial. Other models are natural extensions of this kind of situation with greater choice of primary and secondary distributions and also more levels of modeling by considering models with three (or possibly more) constituent distributions.

When the various constituent distributions; i.e. the number of accidents and the number of claims per accident, can be modeled and estimated separately, compound claim frequency distributions aries naturally. However, they are also useful alternatives when the actuary has difficulty in fitting one of the distributions to claims frequency data.

The moments of compound claim frequency distributions can be evaluated in terms of the moments of the constituent distributions by differentiating the cumulant generating function (cgf)

$$C_N(z) = C_1(C_2(z)) \tag{2.6}$$

and setting $z = 0$. The results for the first four cumulants are:

$$\begin{aligned}
\kappa_1 &= \kappa_{1,1}\kappa_{2,1} \\
\kappa_2 &= \kappa_{1,2}{\kappa_{2,1}}^2 + \kappa_{1,1}\kappa_{2,2} \\
\kappa_3 &= \kappa_{1,3}{\kappa_{2,1}}^3 + 3\kappa_{1,2}\kappa_{2,1}\kappa_{2,2} + \kappa_{1,1}\kappa_{2,3} \\
\kappa_4 &= \kappa_{1,4}{\kappa_{2,1}}^4 + 6\kappa_{1,3}{\kappa_{2,1}}^2\kappa_{2,2} + 3\kappa_{1,2}{\kappa_{2,2}}^2 + 4\kappa_{1,2}\kappa_{2,1}\kappa_{2,3} + \kappa_{1,1}\kappa_{2,4}
\end{aligned} \tag{2.7}$$

Where $\kappa_{i,j}$ is the jth cumulant for the distribution with cgf $C_i(z)$.

An alternative way to compute probabilities is given in the following theorem:

Theorem 2.1: For compound distributions where the primary distribution is a member of the class(a, b) family and where the secondary distribution takes positive values only, the pmf satisfies the following recursion:

$$P[N = n] = \sum_{i=1}^{n} \left(a + \frac{b}{n} i \right) p_2(i) P[N = n - i] \tag{2.8}$$

with the starting value given by $P[N = 0] = p_1(0) = P_1(0)$. Here $p_2(i)$ is the probability function of the secondary distribution.

In the literature one may find two proofs for this theorem; one proof based on generating functions and another proof based on conditional probabilities. Here we present the later proof.

We establish the following lemma to be used in the proof of the theorem.

Lemma 2.1: For $X_1, X_2, ..., X_n$ which are independent and identically distributed random variables taking on values restricted to the positive integers, we have, for positive integer values of x

$$p^{*n}(x) = \sum_{i=1}^{x} p(i) p^{*(n-1)}(x - i) \tag{2.9}$$

$$p^{*n}(x) = \frac{n}{x} \sum_{i=1}^{x} i p(i) p^{*(n-1)}(x - i) \tag{2.10}$$

Proof:

For $n = 1$, both (2.9) and (2.10) reduce to $p^{*1}(x) = p(x) \times p^{(0)}(0)$. For $n > 1$ we establish (2.9) by using the Law of Total Probability to evaluate $Pr(X_1 + X_2 + ... +$

$X_n = x)$ by conditioning on the value taken by X_1 as

$$\sum_{i=1}^{x} Pr(X_1 = i)Pr(X_2 + X_3 + ... + X_n = x - i) \tag{2.11}$$

We then note that $Pr(X_2 + X_3 + ... + X_n = x - i)$ and $Pr(X_1 + X_2 + ... + X_n = x)$ can be evaluated by using $(n-1)$-fold and n-fold convolutions, respectively, of $p(i)$. For $n > 1$, we establish (2.10) by using the conditional expectations $E[X_k|X_1 + X_2 + X_3 + ... + X_n = x]$ for $k = 0, 1, 2, 3, ..., n$. From reasons of symmetry, these quantities are the same for all such k. Since their sum is x, each is equal to x/n. The conditional expectation $E[X_1|X_1 + X_2 + X_3 + ... + X_n = x]$ is evaluated as

$$\sum_{i=0}^{x} i\,\Pr(X_1 = i)Pr(X_2 + X_3 + ... + X_n = x - i)/Pr(X_1 + X_2 + X_3 + ... + X_n = x)$$

We then note that $\Pr(X_2+X_3+...+X_n = x-i)$ and that $\Pr(X_1+X_2+X_3+...+X_n = x)$ can be evaluated by using $(n-1)$-fold and n-fold convolutions, respectively, of $p(i)$. Solving for $p^{*n}(x)$ completes the proof.

Proof of the theorem:

First,

$$\begin{aligned} P[N = 0] &= P[N_1 = 0] \\ P[N = n] &= \sum_{i=1}^{\infty} P[N_1 = i]p_2^{*i}(n), \qquad \text{for } n > 0 \end{aligned}$$

where N_1 is the primary random variable. Since N_1 is a member of class (a, b) distributions

$$\Pr[N_1 = i] = \left(a + \frac{b}{i}\right)\Pr(N_1 = i - 1)\cdot$$

We have

$$P[N = n] = \sum_{i=1}^{\infty} \left(a + \frac{b}{i}\right) p[N_1 = i - 1]p_2^{*i}(n)$$

and by Lemma 2.1

$$\begin{aligned}
P[N=n] &= a\left[\sum_{i=1}^{\infty} P[N_1=i-1]\sum_{j=1}^{n} p_2^{*(i-1)}(n-j)p_2(j)\right] \\
&\quad +b\left[\sum_{i=1}^{\infty}\frac{P[N_1=i-1]}{i}\frac{i}{n}\sum_{j=1}^{n} jp_2(j)p_2^{*(i-1)}(n-j)\right] \\
&= a\left[\sum_{j=1}^{n}\left[\sum_{i=1}^{\infty} P[N_1=i-1]p_2^{*(i-1)}(n-j)\right]p_2(j)\right] \\
&\quad +\frac{b}{n}\left[\sum_{j=1}^{n}\left[\sum_{i=1}^{\infty} P[N_1=i-1]p_2^{*(i-1)}(n-j)\right]p_2(j)\right]
\end{aligned}$$

Hence

$$P[N=n]=\sum_{j=1}^{n} P[N=n-j]\left(a+\frac{b}{n}\right)p_2(j)$$

□

When the distribution of X is Poisson, the resulting distribution is called Neyman Type A distribution. The pgf of the Neyman Type A distribution is

$$P(z)=exp\{\lambda_1[e^{\lambda_2(z-1)}-1]\} \tag{2.12}$$

Using the pgf, it is easy to show that the Neyman Type A distribution has mean

$$\mu'_1=\lambda_1\lambda_2 \tag{2.13}$$

variance

$$\mu_2=\lambda_1\lambda_2(1+\lambda_2) \tag{2.14}$$

skewness

$$\frac{1+3\lambda_2+\lambda_2^2}{(\lambda_1\lambda_2)^{\frac{1}{2}}(1+\lambda_2)^{\frac{3}{2}}} \tag{2.15}$$

kurtosis

$$3 + \frac{1 + 7\lambda_2 + 6\lambda_2^2 + \lambda_2^3}{\lambda_1\lambda_2(1+\lambda_2)^2} \tag{2.16}$$

2.4 Mixture Distributions

The notion of mixing often is a simple and direct interpretation of the physical situation under investigation. For instance, the random variable concerned may be the result of actual mixing of a number of different populations, such as the number of car insurance claims per driver, where the expected number of claims varies with category of driver. Alternatively, the random variable may come from a number of different sources, but the source is known; a mixture rv is then the outcome of ascribing a probability distribution to the possible sources. Sometimes, however, "mixing" is just a mechanism for constructing new distributions for which empirical justification must be sought.

The two important categories of mixtures of discrete distributions are as follows:

1. A k-component finite mixture distribution is formed from k different component distributions with cdf's $F_1(x), F_2(x), ..., F_k(x)$ with mixing weights

   $$\omega_1, \omega_2, ..., \omega_k \qquad \text{where,} \quad \omega_j > 0, \qquad \sum_{j=1}^{k} \omega_j = 1$$

   by taking the weighted average

   $$F(x) = \sum_{j=1}^{k} \omega_i F_j(x) \tag{2.17}$$

   as the cdf of the new(mixture) distribution. This corresponds to the actual mixing of a number of different distributions. In his book, Medgyessy(1977) calls this superposition of distributions. In the theory of insurance, $\omega_j, j = 1, ..., k$ is called the risk function. It follows from (2.17) that if the component distributions are defined on the nonnegative integers with

$$P_j(x) = F_j(x) - F_j(x-1) \tag{2.18}$$

then the mixture distribution is a discrete distribution with pmf

$$Pr[X = x] = \sum_{j=1}^{k} \omega_j P_j(x) \tag{2.19}$$

Example 2.1: The weighted average of two Binomial distributions

$$P[X = x] = \begin{pmatrix} m \\ x \end{pmatrix} {p_1}^x (1-p_1)^{m-x} * \omega_1 + \begin{pmatrix} m \\ x \end{pmatrix} {p_2}^x (1-p_2)^{m-x} * (1-\omega_1)$$

2. A mixture distribution also arises when the cumulative distribution function of a rv depends on the parameters $\theta_1, \theta_2, ..., \theta_m$ (i.e. has the form $F(x|\theta_1, ..., \theta_m)$) and some (or all) of those parameters vary according to a certain joint distribution. The new distribution then has the cumulative distribution function

$$E[F(x|\theta_1, ..., \theta_m)]$$

where the expectation is with respect to the joint distribution of the k parameters that vary. This includes the situation where the source of a random variable is unknowable.

When the points of increase of the mixing distribution are continuous, we will call the outcome a continuous mixture. The cdf is obtained by integration over the mixing parameter Θ; if $H(\Theta)$ is the cdf of Θ, then the mixture distribution has cdf

$$F(x) = \int F(x|\theta) dH(\theta)$$

where integration is over all values taken by Θ. From this equation, the pmf for a mixture of discrete distributions formed using a continuous distribution via Bayes theorem is

$$Pr[X = x] = \int Pr[X = x|\theta]h(\theta)d\theta \tag{2.20}$$

where integration is over all values of θ;the probability density function

$$\frac{h(\theta)Pr[X = x|\theta]}{Pr[X = x]}$$

can be looked upon as a posterior density function for the prior density function $h(\theta)$.

Example 2.2: Binomial distribution mixed over a Beta distribution

We have

$$f(x|\theta) = \begin{pmatrix} n \\ x \end{pmatrix} \theta^x(1-\theta)^{n-x}, \qquad x = 0, 1, 2, ..., n$$

where the pdf of Beta distribution is

$$h(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha}(1-\theta)^{\beta-1}, \qquad 0 < \theta < 1$$

then, by summing over the random variable θ, we obtain

$$f(x) = \begin{pmatrix} n \\ x \end{pmatrix} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+x)}{\Gamma(n+\alpha+\beta)}\Gamma(n+\beta-y), \qquad y = 0, 1, 2, ..., n$$

2.4.1 Mixed poisson distribution

Suppose that any particular risk in the portfolio has a Poisson distribution of claim frequencies with mean $\lambda\theta$ where θ is itself a random variable with the distribution

$U(\theta)$. θ can be interpreted as the expected risk inherent in the given portfolio. Then the (unconditional) distribution of claim frequencies of an individual risk drawn from the portfolio is mixed Poisson.

The Poisson assumption is made as a matter of convenience. For instance, mixed Poisson distributions over Inverse Gaussian have been studied extensively in recent years.

The pgf of the number of claims N is easily seen to be (by conditioning on θ),

$$P_N(z) = \int_0^\infty e^{\lambda\theta(z-1)} dU(\theta)$$

If the Laplace transform of the mixing variate θ is given by

$$L_\theta(z) = \int e^{-z\theta} dU(\theta)$$

then

$$P_N(z) = L_\theta[\lambda(1-z)]$$

The mixed Poisson variates have variance exceeding the mean (unlike the Poisson). This condition, which is usually the case in particular situations, is normally desirable from the insurer's standpoint in that the mixed distribution can be thought of as being "safer" than the original Poisson. Moreover, the convolution of two mixed Poisson variates is again a mixed Poisson variate. Thus, mixed Poisson distributions are closed under convolution.

2.4.2 An example: Poisson mixture with Gamma distribution

For Poisson mixture, we write,

$$P_N(z|\theta) = e^{\lambda(z-1)}$$

then by taking the expectation with respect to the random variable θ, we have

$$P_N(z) = \int_0^\infty e^{\lambda\theta(z-1)} f_\theta(\theta) d(\theta)$$

If we take Gamma distribution with parameter α and β as our mixing distribution, we have

$$P_N(z) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty \theta^{(\alpha-1)} e^{-\theta[\frac{1}{\beta-\lambda(z-1)}]} d\theta.$$

Upon simplification we have

$$P_N(z) = \frac{1}{\beta^\alpha [\frac{1}{\beta-\lambda(z-1)}]^\alpha},$$

or

$$P_N(z) = \frac{1}{[1-\lambda\beta(z-1)]^\alpha}.$$

This is the pgf of a Negative Binomial distribution with parameters $\lambda\beta$ and α.

2.4.3 Generalized Poisson-Pascal Distribution

This three-parameter distribution has a Poisson primary distribution and a secondary distribution that is a truncated negative binomial or an extended truncated negative binomial. Consequently, its pgf is

$$P(z) = \exp\left[\lambda \frac{[1-\beta(z-1)]^{-r} - (1+\beta)^{-r}}{1-(1+\beta)^{-r}} - 1\right], \qquad r > -1, \lambda, \beta > 0. \qquad (2.21)$$

In the case when $r > 0$,this pgf is that of the Poisson-Pascal distribution with a Poisson primary distribution with mean $\lambda[1-(1+\beta)^{-r}]^{-1}$ and a negative binomial secondary distribution. When $r = 1$ this distribution is known as the Polya-Aeppli. When $r = -.5$ the distribution is known as Poisson-Inverse Gaussian distribution. The Poisson-Pascal distribution is generalized by extending the range of r to -1.

It is necessary to consider the primary distribution to be Poisson with mean λ and the secondary distribution to be truncated at zero. This would lead to recursively computable probabilities.

The probabilities of the generalized Poisson-Pascal distribution can be calculated recursively as

$$p_n = \frac{\lambda}{n} \sum_{j=1}^{n} j q_j p_{n-j}, \qquad n = 1, 2, ..., \tag{2.22}$$

where

$$p_0 = e^{-\lambda} \tag{2.23}$$

and $\{q_n : n = 1, 2, ..\}$ are probabilities from the (extended) truncated negative binomial distribution which can be calculated recursively as

$$q_n = \frac{n+r-1}{n} \left(\frac{\beta}{1+\beta} \right) q_{n-1}, \qquad n = 2, 3... \tag{2.24}$$

beginning with

$$q_1 = \frac{r}{(1+\beta)^r - 1} \left(\frac{\beta}{1+\beta} \right) \tag{2.25}$$

2.5 The $(a+b, b)$ class of distributions

Consul and Jain (1973) and Consul and Shoukri (1985) introduced the generalized Poisson distribution (GPD) with the probability function:

$$P[N = n] = \frac{\lambda(\lambda + n\theta)^{n-1}}{n!} \exp(\lambda - n\theta), \qquad n = 0, 1, ..., \quad \lambda > 0, \quad 0 \leqslant \theta < 1 \tag{2.26}$$

Consul (1990) compared this distribution with a number of other distributions with respect to fitting claims counts and concluded that it is a plausible model.

Ambagaspitiya and Balakrishnan (1994) observed that the generalized Poisson distribution (GPD) satisfies the following recursion

$$p_n(\lambda,\theta) = \frac{\lambda}{\lambda+\theta}\left(\theta+\frac{\lambda}{n}\right)p_{n-1}(\lambda+\theta,\theta) \qquad n=1,2,... \tag{2.27}$$

and illustrated that GPD yeld recursively computable total claims distributions.

Ambagaspitiya (1995) showed that there are large numbers of distributions satisfying the following generalized version of the recursion (2.27)

$$p_n(a,b) = \left(h_1(a,b)+\frac{h_2(a,b)}{n}\right)p_{n-1}(a+b,b), \qquad n=1,2,... \tag{2.28}$$

Where $h_1(a,b), h_2(a,b)$ are two functions of the parameters a and b.

The following list gives a number of important distributions in this class.

1. Generalized Poisson distribution.

2. Generalized negative binomial distribution, presented by Consul and Gupta (1980), with the probability function

$$p_n(a,b) = \frac{a}{a+bn}\begin{pmatrix} a+bn \\ n \end{pmatrix}\alpha^n(1-\alpha)^{a+bn-n}, \quad a>0, \alpha>0, 1\leqslant b<\frac{1}{\alpha} \tag{2.29}$$

3. Discrete distributions obtained by weighted Generalized Poisson or generalized negative binomial distributions. The term weighting stands for multiplying the probability $p_n(a,b)$ by a function of the form $w(a+bn;b)$ for each $n=0,1,2,...$ and then dividing by the normalizing constant.

Chapter 3

Models for Claim Sizes

3.1 Introduction

In an insurance portfolio, the smallest possible claim size would be one cent. Therefore, we could represent claim amounts by discrete random variables. However, this would lead to an unmanageable support for portfolios with potentially large claims such as in liability insurance. Therefore practitioners use continuous random variables to represent claim sizes.

In this chapter we present four distributions that can be used in modeling claim sizes. Two of them are well known in the statistical literature but the other two are somewhat obscure.

3.2 Log-normal Distribution

The probability density function of log-normal distribution is

$$f(x) = \frac{1}{\sigma x \sqrt{(2\pi)}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right], \qquad x > 0, \quad \sigma > 0 \tag{3.1}$$

and the cumulative distribution function is

$$F(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right). \tag{3.2}$$

Where $\Phi(\cdot)$ is the cumulative distribution function of standard normal variable. It can easily be seen that the lognormal random variable is obtained by using the transformation $X = e^Y$, where Y is normal random variable with mean μ and variance

σ^2.

A three parameter lognormal distribution can be obtained by introducing a location parameter θ:

$$f(x) = \frac{1}{\sigma(x-\theta)\sqrt{(2\pi)}} \exp\left[-\frac{1}{2}\left(\frac{\ln(x-\theta)-\mu}{\sigma}\right)^2\right], \quad x > \theta, \sigma > 0 \cdot \tag{3.3}$$

The lognormal distribution is sometime called the antilognormal distribution. This name has some logical basis in that it is not the distribution of the logarithm of a normal variable (this is not even always real) but of an exponential - that is, antilogrithmic - function of such variable. However, "Lognormal" is most commonly used. The minor variants logarithmic - or logarithmico-normal have been used, as have the names of pioneers in its development, notably Galton(1879) and McAlister (1879), Kapteyn(1903), van Uven(1917a) and Gilbrat (1930). When applied to economic data, particularly production functions, it is sometimes called the Cobb-Douglas distribution.

The important case ($\theta = 0$) has been given the name two-parameter lognormal distribution (parameters μ and σ). The rth moment of lognormal (two-parameter) distribution is

$$\mu_r' = E[X^r] = E[e^{Yr}] = \exp\left(\mu r + \frac{1}{2}\sigma^2 r^2\right) \tag{3.4}$$

The expected value of X is

$$\mu_1' = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

and variance of X is

$$\mu_2 = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1]$$

We can estimate the parameter by taking log of the data and using maximum likelihood methods to get the parameter, i.e.

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \log x_i = \log\left[\prod_{i=1}^{n} x_i^{1/n}\right]$$

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n} (\log x_i - \hat{\mu})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\log x_i - \log\left[\prod_{i=1}^{n} x_i{}^{1/n}\right]\right]^2 \\
&= \log\left[\prod_{i=1}^{n}\left[\frac{x_i}{\prod_{j=1}^{n} x_j^{1/n}}\right]^{1/n}\right]
\end{aligned}$$

An additional property of the lognormal distribution is that if $X_1, X_2, ..., X_n$ is a set of independently and identically distributed lognormal random variables such that the mean of each $\log X_i$ is μ and its variance is σ^2 , then the product $X_1X_2X_3...X_n$ is distributed as lognormal with mean and variance of $\log(X_1X_2...X_n)$ as $n\mu$ and $n\sigma^2$ respectively.

3.3 Pareto Distribution

The Pareto distribution has probability density function

$$f(x) = \frac{\alpha\beta^{\alpha+1}}{x^{\alpha+1}}, \qquad x > \beta \tag{3.5}$$

and distribution function

$$F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}, \qquad x > \beta \tag{3.6}$$

The Pareto distribution is named after a Swiss professor of economics, Vilfredo Pareto(1848-1923). Pareto's law, as formulated by him(1897),dealt with the distribution of income over a population and can be stated as follows:

$$N = Ax^{-\alpha}$$

where N is the number of persons having income $\geqslant 0$, and A, α are parameters (α is known both as pareto's constant and as a shape parameter). It was felt by Pareto that this law was universal and inevitable - regardless of taxation and social and political conditions. " Refutations" of the law have been made by several well-known economists over the past 60 years [e.g.,Pigou (1932); Shirras (1935); Hayakawa (1951)]. More recently attempts have been made to explain many empirical phenomena using the Pareto distribution or some closely related form[e.g., Steindl (1965); Mandelbrot (1960,1963,1967); Hagstroem (1960); Ord (1975)].

Harris (1968) has pointed out that a mixture of exponential distributions, with parameter θ^{-1} having a gamma distribution, and with origin at zero, gives rise to a Pareto distribution [Maguire, Pearson, and Wynn (1952)]. In fact, if

$$Pr[X \leqslant x|\theta] = 1 - e^{-x/\theta}$$

and $\mu = \theta^{-1}$ has a gamma distribution, then

$$\begin{aligned} Pr[X \leqslant x] &= \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_0^{\infty} t^{\alpha-1} e^{-t/\beta}(1 - e^{-tx})dt \\ &= 1 - (\beta x + 1)^{-\alpha}, \qquad x \geqslant 0 \end{aligned} \tag{3.7}$$

which is the form of the Pareto distribution of the second kind.

The pdf of the Pareto distribution(3.5)is the special form of a Pearson Type VI

distribution; we denoted it by $X \sim P(I)(\beta, \alpha)$. The relation given by (3.6) is now more properly known as the pareto distribution of the first kind.

The rth ($r < \alpha$) moment about zero of a Pareto distribution is

$$\begin{aligned} \mu'_r &= \alpha\beta^r(\alpha - r)^{-1} \\ &= \frac{\alpha\beta^r}{\alpha - r}, \qquad r < \alpha \cdot \end{aligned} \tag{3.8}$$

In particular the expected value is

$$\begin{aligned} E[X] &= \alpha\beta(\alpha - 1)^{-1} \\ &= \frac{\alpha\beta}{\alpha - 1}, \qquad \alpha > 1, \end{aligned} \tag{3.9}$$

and the variance is

$$\begin{aligned} Var[X] &= \alpha\beta^2(\alpha - 1)^{-1}(\alpha - 2)^{-1} \\ &= \frac{\alpha\beta^2}{(\alpha - 1)(\beta - 1)}, \qquad \alpha > 2 \cdot \end{aligned} \tag{3.10}$$

In actuarial literature the Pareto distribution function takes the form

$$F(y) = 1 - \left(\frac{\lambda}{\lambda + y}\right)^{\alpha}, \qquad y \geqslant 0, \quad \lambda > 0, \alpha > 0 \cdot \tag{3.11}$$

We see that this distribution can be obtained from the original Pareto distribution by using the transformation $Y = X - \beta$ and then writing λ for β. Therefore, it is the Pareto distribution of second kind with $\beta = 1/\lambda$.

Also, in the actuarial literature the distribution obtained by the transformation $Z = (X - \beta)^{1/\tau}, \tau > 0$ and $\lambda = \beta$ where the distribution of X is pareto, is called the Burr distribution.

The cdf of it takes the form

$$F(z) = 1 - \left(\frac{\lambda}{\lambda + z^{\tau}}\right)^{\alpha} \tag{3.12}$$

Note that Burr (1942) presented 12 distributions and the 12th distribution has the cdf

$$F(y) = 1 - \left(\frac{1}{1+y^c}\right)^k \tag{3.13}$$

If we use the transformation $Z = aY$ and $\lambda = a^c; \alpha = k$ we can get the Burr distribution in actuarial literature. Moments of the Burr distribution are given by

$$E[X^n] = \lambda^{n/\tau}\left(\alpha - \frac{n}{\tau}\right)\frac{\Gamma(1+n/\tau)}{\Gamma(\alpha)}, \qquad \alpha\tau > n,\ n = 1, 2, ... \tag{3.14}$$

3.4 Gamma Distribution

The probability density function of gamma distribution is in the form:

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-x/\beta\right), \qquad x > 0, \quad \alpha, \beta > 0, \tag{3.15}$$

where

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \cdot$$

The main application of the Gamma distribution in actuarial literature is using it as the prior distribution of the Poisson paremeter.

By introducting a location shift we could obtain the following three parameter gamma distribution.

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha}(x - x_0)^{\alpha-1} \exp\left[-(x - x_0)/\beta\right], \qquad x > x_0, \quad \alpha, \beta > 0 \ \cdot \tag{3.16}$$

This distribution is known as the translated gamma distribution and, as we explain in Chapter 4 actuaries, use this to approximate compound distributions.

When the parameter $\alpha = n$ is an integer, it is called Erlang's distribution. In this case the cumulative distribution function can be expressed as

$$F(x) = 1 - \sum_{i=1}^{n} \frac{(x/\beta)^i}{i!} \exp\left(-x/\beta\right) \tag{3.17}$$

Also by setting $\alpha = r/2$ and $\beta = 2$, where r is an integer, we obtain the well known chi-square distribution. The μ_r', rth non-central moment, of the gamma distribution can be written as

$$\mu_r' = \frac{\Gamma(\alpha + \beta^r)}{\Gamma(\alpha + r)} \beta^r \,. \tag{3.18}$$

3.5 Weibull Distribution

The probability density function of the Weibull distribution is

$$f(x) = c\tau x^{(\tau-1)} e^{-cx^\tau} \tag{3.19}$$

and the distribution function is

$$F(x) = 1 - e^{-cx^\tau} \tag{3.20}$$

The Weibull distribution is named after the Swedish physicist, Waloddi Weibull, who (1939a,b) used it to represent the distribution of the breaking strength of materials and, in 1951, for a wide variety of other applications. In the Russian statistical literature this distribution is often referred to as the Weibull-Gnedenko distribution. The Weibull distribution includes the exponential and the Rayleigh distributions as special cases.

The rth non-central moment of the Weibull distribution is

$$E[X^r] = \frac{\Gamma\left(1 + r/\tau\right)}{c^{r/\tau}} \tag{3.21}$$

Chapter 4

Computing Aggregate Claims Distribution

4.1 Introduction

The main reason behind the analysis of claims data is revising the premiums for a portfolio. In the early days of actuarial practice the pure premium was set as the mean of the total claims plus a security loading. This security loading was a factor of the standard deviation of total claims. Then the pure premium was loaded with expenses to compute gross premiums. This method did not take long tail nature of the claims distribution. At present actuaries set pure premium as percentiles of the total claims distribution. This is due to advances in the computation of total claims distributions instead of just means and variances.

In this chapter we review available methods to compute total claims distribution and discuss their advantages and disadvantages. Although, the discussion is general attention is focused on computing total claims distribution when claim counts are generalized Poisson. Section 2 describes the exact computation. Section 3 describes the moment approximation and Section 4 describes the inverting generating functions and in Section 5, we present the recursive method. Finally, in Section 6, we describe how to use the recursive method with non-arithmetic severities. We implement this method with VBA.

4.2 Direct Computation

To derive the d.f. of S we distinguish according to how many claims occur and use the law of total probability

$$F_S(x) = Pr(S \leqslant x) = \sum_{n=0}^{\infty} Pr(S \leqslant x | N = n) Pr(N = n)$$
$$= \sum_{n=0}^{\infty} Pr(X_1 + X_2 + ... + X_n \leqslant x) Pr(N = n) \tag{4.1}$$

But,in terms of convolution operation, we can write

$$Pr\,(X_1 + X_2 + ... + X_n \leqslant x) = F * F * F * ... * F(x)$$
$$= F^{*n}(x) \tag{4.2}$$

called the nth convolution of F. Thus (4.1) becomes

$$F_S(x) = \sum_{n=0}^{\infty} F^{*n}(x) Pr(N = n) \tag{4.3}$$

If the individual claim amount distribution is discrete with a p.f. $f(x) = Pr(X = x)$, the distribution of aggregate claims is also discrete. By analogy with the distribution, the p.f. of S can be obtained directly as

$$f_S(x) = \sum_{n=0}^{\infty} f^{*n}(x) Pr(N = n) \tag{4.4}$$

where

$$f^{*n}(x) = f * f * .... * f(x) = Pr(X_1 + X_2 + ... + X_n = x) \tag{4.5}$$

Here, the inequality sign in the probability symbol in (4.1) has been replaced by the equality sign.

Example 4.1: Consider an insurance portfolio that will produce $0, 1, 2$, or 3 claims in a fixed time period with probabilities $0.1, 0.3, 0.4$ and 0.2 respectively. An individual claim amount $1, 2$, or 3 with probability $0.5, 0.4, 0.1$ respectively. For this portfolio we illustrate the direct computation of total claims distributions. Since there are 3

Table 4.1: Direct computation

x	$p^{*0}(x)$	$p^{*1}(x) = p(x)$	$p^{*2}(x)$	$p^{*3}(x)$	$f(x)$	$F(x)$
0	1.0	–	–	–	0.1000	0.1000
1	–	0.5	–	–	0.1500	0.2500
2	–	0.4	0.25	–	0.2200	0.4700
3	–	0.1	0.40	0.125	0.2150	0.6850
4	–	–	0.26	0.300	0.1640	0.8490
5	–	–	0.08	0.315	0.0950	0.9440
6	–	–	0.01	0.184	0.0408	0.9848
7	–	–	–	0.063	0.0126	0.9974
8	–	–	–	0.012	0.0024	0.9998
9	–	–	–	0.001	0.0002	1.0000
n	0	1	2	3	–	–
$Pr(N = n)$	0.1	0.3	0.4	0.2	–	–

claims and each produces a claim amount of at most 3, we can limit the calculations to $x = 0, 1, 2, ...9$. Column(2) lists the p.f. of a degenerate distribution with all the probability mass at 0. Column (3) lists the p.f. of the individual claim amount random variable. Column (4) and (5) are obtained recursively by applying

$$\begin{aligned} f^{*(n+1)}(x) &= Pr(X_1 + X_2 + ... + X_{n+1} = x) \\ &= \sum_y Pr(X_{n+1} = y) Pr(X_1 + X_2 + ... + X_{n+1} = x - y) \\ &= \sum_y f(y) f^{*n}(x - y) \qquad (4.6) \end{aligned}$$

Since only 3 different claims amounts are possible, the evaluation of (4.6) will involve a sum of 3 or fewer terms. Next, (4.4) is used to compute the p.f. displayed in column(6). For this step, it is convenient to record the p.f. of N in the last row of the results. Finally, the elements of column (7) are obtained as partial sums of column (6). An alternative approach, would have been to perform the convolutions in terms of the d.f.'s, obtain $F(x)$ from (4.3), and finally $f(x) = F(x) - F(x-1)$.

4.3 Moment Approximations

The most used moment approximation in the actuarial literature is the normal approximation. Normal approximation is equating the first two central moment to the mean and variance of a normal distribution.

Example 4.2: Let us assume the claim count distribution is Generalized Poisson with parameters ($\lambda = 10,\ \theta = 0.2$) and the claim size distribution is gamma with parameters ($\alpha = 2,\ \ \beta = 0.5$). Since the mean and the variance of the compound generalized distribution takes the form

$$\begin{aligned} E[S] &= \lambda p_1 M \\ Var[S] &= \lambda p_2 M^3 + \lambda(p_2 - p_1^2)M \end{aligned}$$

where $M = (1-\theta)^{-1}$ and p_1 and p_2 are first and second non-central moments of claim sizes distribution, we have the

$$\begin{aligned} E[S] &= \frac{10}{1-0.2} * \frac{2}{0.5} \\ &= 50\cdot \\ Var[S] &= 12.5 * \frac{2}{0.5^2} + \left(\frac{2}{0.5}\right)^2 * \frac{10}{(1-0.2)^3} \\ &= 412.5\cdot \end{aligned}$$

Therefore, the normal approximation says that the total claims distribution is normal with mean 50 and variance 412.5. This gives us the probability of having negative total claims is about 0.6%. Although, this is an extreme example, it is obvious that the normal approximation is not appropriate in many cases.

Since the normal approximation yields undesirable results, three moment approximations has been studied in the actuarial literature. Although, any distribution with three parameters can be used as the approximated distribution, translated gamma distribution is the most promising distribution. Bowers et al. (1997) credit Seal (1978a) as the originator of this method.

The translated gamma distribution is the gamma with a location shift. Its probability density function takes the form

$$f(x|\alpha,\beta,x_0) = \frac{\beta^\alpha}{\Gamma(\alpha)}(x-x_0)^{\alpha-1}\exp(-\beta(x-x_0)), \quad x \geq x_0 \cdot \tag{4.7}$$

The parameters x_0, α and β are obtained by solving the following equations

$$\begin{aligned} E[S] &= x_0 + \frac{\alpha}{\beta}, \\ Var[S] &= \frac{\alpha}{\beta^2}, \\ E[(S-E[S])^3] &= \frac{2\alpha}{\beta^3} \cdot \end{aligned}$$

Let us look at the previous example again. Since the third central moments of the compound generalized distribution takes the form:

$$E[(S-E[S])^3] = \lambda(3M-2)p_1^3M^4 + 3\lambda p_1(p_2-p_1)^2M^3 + (p_3 - 3p_2p_1 + 2p_1^3)\lambda M$$

where $M = (1-\theta)^{-1}$ and p_3 is the 3rd non-central moments of claim severity. For our example $M = 1.25$ and $p_1 = 4, p_2 = 24, p_3 = 192$ we have the third central moment of S as 5009.375. Therefore we have three equations for parameters

$$\begin{aligned} x_0 + \frac{\alpha}{\beta} &= 50 \\ \frac{\alpha}{\beta^2} &= 412.5 \\ \frac{2\alpha}{\beta^3} &= 5009.375 \end{aligned}$$

Solving these three equations we have

$$\begin{aligned} x_0 &= -17.935 \\ \alpha &= 11.1883 \\ \beta &= 0.1647 \end{aligned}$$

Therefore the approximation indicates that the pdf of S takes the form

$$f_S(s) = \frac{0.1647^{11.1883}}{\Gamma(11.1883)}(x - x_0)^{10.1883} \exp(-0.1647(x + 17.935)), \quad x > -17.935$$

A further generalization of the moment approximation would be to compute the first four central moments of the total claim distribution and then fitting a distribution in the Pearson system.

Johnson, Balakrishnan and Kotz (1994) present the Pearson system as the system satisfying the differential equation

$$\frac{1}{f(x)}\frac{df(x)}{dx} = -\frac{a + x}{c_0 + c_1 x + c_2 x^2}.$$

This family includes a large number of continuous distributions. This method has been explored by Kaas, Goovaerts (1985). However, we were unable to find any follow up work of this method in the actuarial literature.

4.4 Inverting generating functions

In most cases it is easy to obtain generating functions. Therefore actuarial researchers have used this method. If the claim distribution is discrete then we can compute the probability function of the total claims distribution through probability generating functions. This is best illustrated through an example.

Example 4.3: Suppose that S has a compound generalized Poisson distribution with $\lambda = 4/5$ and $\theta = 1/2$ and the individual claim amount distribution is as follows: Then the probability generating function of the claim sizes is

Table 4.2: Claim size distribution

x	1	2	3
$\Pr(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$$P_x(z) = \frac{z}{4} + \frac{z^2}{2} + \frac{z^3}{4}.$$

Then the probability generating function of S takes the form

$$P_S(z) = \exp\left(-\frac{8}{5}W(-\frac{1}{2e^{1/2}}P_x(z)) - \frac{4}{5}\right),$$

where W(x) is the inverse function of xe^x or $W(x)e^{W(x)} = x$.

We could use Maple to find the series expansion of this up to 5 terms to obtain

$$P_S(z) = 0.44933 + 0.054508z + 0.11646z^2 + 0.085367z^3 + 0.051433z^4 + 0.0477146z^5 + O(z^6)\cdot$$

From this we could obtain the following probabilities:

Table 4.3: Total claim distribution

s	0	1	2	3	4	5
$\Pr[S=s]$	0.44933	0.054508	0.11646	0.085367	0.051433	0.0477146

In theory we could use this method to compute $\Pr[S = s]$ for any value of s; however there are a few problems associated with this method. In realistic situations, the individual claim sizes variable could take a large number of values resulting in cumbersome generating function. Also, packages such as Maple and Mathematics are not very popular among practising actuaries.

4.4.1 Continuous individual claim sizes distribution

In this situation we could invert the characteristic function using the inversion theorem. Inversion theorem states that the characteristic function $C(t)$

$$C(t) = \int_{-\infty}^{\infty} \exp(itx) f(x) dx ,$$

uniquely determines the pdf $f(x)$ via

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx) C(t) dt\cdot$$

The proof of this theorem is given in Stuart and Ord (1994, p.126). Since the total claim distribution is of mixed type we have

$$C(t) = f_S(0) + \int_{0}^{\infty} \exp(itx) f_S(x) dx ,$$

with $f_S(0) = \Pr[N = 0]$. Therefore

$$f_S(x) = \int_{-\infty}^{\infty} (C(t) - f_S(0)) \exp(-itx) dt. \tag{4.8}$$

In most cases the integral in (4.8) has to be evaluated numerically. Heckman and Meyers (1983) illustrate this method for class (a,b) compound distributions. At present, Maple can be used to compute (4.8) for this class without using any of the details presented in the paper. However Maple is too slow for the compound generalized Poisson case.

4.5 Recursive Methods

Since the Panjer(1981) and Sundt and Jewell(1981) rediscovery of recursive computations of compound distributions many actuaries have investigated recursive computations for various compound distributions. Ambagaspitiya(1994) presents a somewhat different recursive method for computing the compound generalized Poisson distribution. The folowing theorem states this method.

Theorem 4.2: If the claim sizes are random variables on the positive integers with probability mass function $f(x) = \Pr(X = x)$, $x = 1, 2, \ldots$, then the probability mass function $g(a, b; x)$ of compound generalized Poisson distribution satisfies the recurrence equation

$$g(a, b; x) = \frac{a}{a+b} \sum_{y=1}^{x} \left(b + a\frac{y}{x}\right) g(a+b, b; x-y) f(y)$$

Proof:

Since

$$p_i(a, b) = \frac{a}{a+b} \left(b + \frac{a}{i}\right) p_{i-1}(a+b, b), \quad i = 1, 2, \ldots$$

and

$$g(a, b; 0) = p_0(a, b) \tag{4.9}$$

$$g(a,b;x) = \sum_{i=1}^{\infty} p_i(a,b;x) f^{*i}(x), \qquad x > 0 \tag{4.10}$$

we have

$$\sum_{i=1}^{\infty} p_i(a,b) f^{*i}(x) = \frac{a}{a+b}\left(b \sum_{i=1}^{\infty} p_{i-1}(a+b,b) f^{*i}(x) + a \sum_{i=1}^{\infty} \frac{p_{i-1}(a+b,b)}{i} f^{*i}(x)\right)$$

Using the identities,

$$f^{*i}(x) = \sum_{y=1}^{x} f^{*i-1}(x-y) f(y) \tag{4.11}$$

$$\frac{f^{*i}(x)}{i} = \sum_{y=1}^{x} \frac{y}{x} f^{*i-1}(x-y) f(y) \tag{4.12}$$

we have

$$\begin{aligned} \sum_{i=1}^{\infty} p_i(a,b) f^{*i}(x) &= \frac{a}{a+b}\left(b \sum_{i=1}^{\infty} p_{i-1}(a+b,b) \sum_{y=1}^{x} f^{*i-1}(x-y) f(y)\right) \\ &\quad + \frac{a}{a+b}\left(a \sum_{i=1}^{\infty} p_{i-1}(a+b,b) \sum_{y=1}^{x} \frac{y}{x} f^{*i-1}(x-y) f(y)\right). \end{aligned}$$

By interchanging the order of summation

$$\sum_{i=1}^{\infty} p_i(a,b) f^{*i}(x) = \frac{a}{a+b}\left[\sum_{y=1}^{x}\left(\sum_{i=1}^{\infty} p_{i-1}(a+b,b) f^{*(i-1)}(x-y)\right)\left(b + \frac{ay}{x}\right) f(y)\right].$$

Since

$$\begin{aligned} \sum_{i=1}^{\infty} p_{i-1}(a+b,b) f^{*i-1}(x-y) &= \sum_{i=0}^{\infty} p_i(a+b,b) f^{*i}(x-y)) \\ &= g(a+b,b;x-y), \end{aligned} \tag{4.13}$$

we obtain

$$\sum_{i=1}^{\infty} p_i(a,b) f^{*i}(x) = \frac{a}{a+b} \sum_{y=1}^{x} \left(b + a\frac{y}{x}\right) g(a+b,b;x-y) f(y)$$

which completes the proof of the theorem. □

4.6 Recursive Computation with Non-Arithmetic Claim Sizes

Ambagaspitiya and Balakrishnan (1994) provides the following integral equation for the compound generalized distribution when claim sizes are continuous.

$$g(a,b;x) = p_1(a,b)f(x) + \frac{a}{a+b}\int_0^x \left(b + a\frac{y}{x}\right) g(a+b,b;x-y)f(y)dy \qquad (4.14)$$

Solving this integral equation is difficult; numerical techniques have to be employed. However, the straightforward method is discretizing the continuous claim sizes distribution and then applying the recurrence scheme outline in Section 4.5. In the literature one may find a number of discretization techniques(see, for example, Panjer and Willmot (1992) pages 223-230). However, we used the method of rounding to the upper unit. In this method, we construct discrete probability distribution with the probability mass at the point x as

$$p(x) = P[x-1 < X \leqslant x] \qquad (4.15)$$

$$p(x) = F_X(x) - F_X(x-1), \qquad x = 1,2,..., \qquad (4.16)$$

where $F_X(x)$ is the cumulative distribution function of the claim sizes (continuous) random variable. In (4.16), x is expressed in terms of smallest monetary unit(for example cents). However, if the claim sizes could go upto millions of dollars we may not need to compute probabilities of total claims to the last cent. Therefore we may use the span of h; that is, we may write

$$p(x) = F_X(h*x) - F_X(h*(x-1)), \qquad x = 1,2,..., \qquad (4.17)$$

h could take any positive integer value. For example $h = 100$ means our calculations are in $100s$.

4.6.1 Implementation in Visual Basic for Applications

In visual basic we could write recursive functions and procedures. However, we found recursive functions tend to be very slow. Therefore, we implemented the recursive procedure using arrays as described in Ambagaspitiya and Balakrishan (1994). This method for computing $g(a,b;x),\ x=0,1,2,3,4$ can be summarized in the following scheme:

$$\begin{array}{lllll}
g(a,b;0) & g(a+b,b;0) & g(a+2b,b;0) & g(a+3b,b;0) & g(a+4b,b;0) \\
g(a,b;1) & g(a+b,b;1) & g(a+2b,b;1) & g(a+3b,b;1) & \\
g(a,b;2) & g(a+b,b;2) & g(a+2b,b;2) & & \\
g(a,b;3) & g(a+b,b;3) & & & \\
g(a,b;4) & & & &
\end{array}$$

The first row of the above scheme is obtained by using the fact that $g(a+ib,b;0) = p_0(a+ib,b) = \exp(-a-ib)$ for $i=1,2,....$ To calculate the probability mass function given in the (i,j)th location, one has to use the elements in $(l,j+1)$ where $l=0,1,...,i-1$. Since the scheme is of an upper diagonal form, we can carry out the computations for each row starting from either left to right or right to left.

In the program we developed 3 forms to interact with the user. The first form lets the user choose parameters of the generalized Poisson distribution. After a valid entry, the user will be directed to the second form where he can choose a distribution from a list of distributions for claim sizes. The third form lets the user choose parameters of the chosen claim size distribution and the span. Then it performs the calculation and the results will be stored in an Excel workseet. Some sample outputs are given in Table 4.4 and Table 4.5.

For the Burr and Weibull distributions we have analytical forms for the cdf so we could use them in discretization. However, for lognormal and gamma distributions we need to use the functions in the Excel library.

Table 4.4: Generalized Poisson distribution with $\lambda = 0.8$ and $\theta = 0.5$; span=1

Lognormal ($\mu = 2, \sigma = 0.5$)		Gamma ($\alpha = 3.5, \beta = 2.7$)		Weibull ($c = 0.62, \tau = 1.3$)		Burr ($\alpha = 3.8$, $\lambda = 1.7, \tau = 0.8$)	
s	$P(S = s)$	s	$P(S = s)$	s	$P(S = s)$	s	$P(S = s)$
0	0.449329	0	0.449329	0	0.449329	0	0.449329
1	6.91E-06	1	0.000435	1	0.100740	1	0.180439
2	0.000969	2	0.003284	2	0.095325	2	0.104149
3	0.00681	3	0.007948	3	0.073195	3	0.067862
4	0.016163	4	0.012715	4	0.055742	4	0.047029
5	0.023478	5	0.016555	5	0.042957	5	0.033874
6	0.026612	6	0.019111	6	0.033556	6	0.025066
7	0.02648	7	0.020444	7	0.026527	7	0.018928
8	0.024737	8	0.020808	8	0.021184	8	0.014521
9	0.022609	9	0.020496	9	0.017068	9	0.011285
10	0.020689	10	0.019771	10	0.013859	10	0.008864
11	0.01912	11	0.018830	11	0.011330	11	0.007027
12	0.017845	12	0.017809	12	0.009320	12	0.005614
13	0.01676	13	0.016791	13	0.007707	13	0.004516
14	0.015785	14	0.015819	14	0.006405	14	0.003655
15	0.014879	15	0.014913	15	0.005346	15	0.002974
16	0.014028	16	0.014076	16	0.004479	16	0.002433
17	0.013228	17	0.013306	17	0.003767	17	0.001998
18	0.012479	18	0.012595	18	0.003178	18	0.001648
19	0.011781	19	0.011937	19	0.002690	19	0.001365
20	0.011131	20	0.011326	20	0.002283	20	0.001134
21	0.010526	21	0.010755	21	0.001942	21	0.000945
22	0.009961	22	0.010221	22	0.001656	22	0.00079
23	0.009433	23	0.009720	23	0.001415	23	0.000663
24	0.008939	24	0.009248	24	0.001212	24	0.000558
25	0.008476	25	0.008805	25	0.001040	25	0.000470
26	0.008041	26	0.008387	26	0.000894	26	0.000398
27	0.007634	27	0.007992	27	0.000769	27	0.000337
28	0.007251	28	0.007620	28	0.000663	28	0.000286
29	0.006891	29	0.007269	29	0.000573	29	0.000244
30	0.006553	30	0.006937	30	0.000495	30	0.000208
31	0.006234	31	0.006624	31	0.000429	31	0.000178
32	0.005934	32	0.006327	32	0.000372	32	0.000152
33	0.005651	33	0.006046	33	0.000323	33	0.000131
34	0.005384	34	0.005780	34	0.000280	34	0.000113
35	0.005131	35	0.005528			35	9.72E-05
36	0.004893	36	0.005289			36	8.41E-05
37	0.004668	37	0.005063			37	7.28E-05
38	0.004455	38	0.004847				
39	0.004253	39	0.004643				
40	0.004062	40	0.004449				
41	0.003881	41	0.004265				
42	0.003710						
43	0.003547						
44	0.003393						
45	0.003246						
46	0.003107						
47	0.002975						
48	0.002849						

Table 4.5: Generalized Poisson Distribution with $\lambda = 0.9$ and $\theta = 0.6$; span=100

Lognormal ($\mu = 4.2, \sigma = 0.8$)		Gamma ($\alpha = 10, \beta = 100$)		Weibull ($c = 0.0025, \tau = 1.0$)		Burr ($\alpha = 2.25, \lambda = 5, \tau = 0.7$)	
s	$P(S = s)$	s	$P(S = s)$	s	$P(S = s)$	s	$P(S = s)$
0	0.40657	0	0.40657	0	0.406570	0	0.40657
1	0.139314	1	2.24E-08	1	0.042849	1	0.197285
2	0.100148	2	9.32E-06	2	0.038975	2	0.113843
3	0.071088	3	0.000212	3	0.035517	3	0.072724
4	0.052989	4	0.001412	4	0.032425	4	0.049545
5	0.040609	5	0.004759	5	0.029655	5	0.035277
6	0.031794	6	0.010462	6	0.027168	6	0.025938
7	0.025315	7	0.017188	7	0.024932	7	0.019544
8	0.020433	8	0.022879	8	0.022918	8	0.015012
9	0.016682	9	0.026001	9	0.021100	9	0.011712
10	0.013752	10	0.026173	10	0.019457	10	0.009256
11	0.011431	11	0.024034	11	0.017968	11	0.007394
12	0.009571	12	0.020706	12	0.016618	12	0.005961
13	0.008065	13	0.017279	13	0.015391	13	0.004844
14	0.006812	14	0.014501	14	0.014274	14	0.003964
		15	0.012689	15	0.013256	15	0.003264
		16	0.011799	16	0.012326	16	0.002702
		17	0.011567	17	0.011475	17	0.002248
		18	0.011654	18	0.010696	18	0.001879
		19	0.011767	19	0.009981	19	0.001576
		20	0.011717	20	0.009325	20	0.001327
		21	0.011433	21	0.008721	21	0.001122
		22	0.010942	22	0.008165	22	0.000951
		23	0.010326	23	0.007651	23	0.000808
		24	0.009679	24	0.007177	24	0.000689
		25	0.009080	25	0.006739	25	0.000589
		26	0.008573	26	0.006333	26	0.000505
		27	0.008167	27	0.005957	27	0.000433
				28	0.005607	28	0.000373
				29	0.005283	29	0.000322
				30	0.004981	30	0.000278
				31	0.004700	31	0.000241
				32	0.004438	32	0.000209
				33	0.004194	33	0.000181
				34	0.003965	34	0.000158
				35	0.003752	35	8.35E-05
				36	0.003552		
				37	0.003365		
				38	0.003190		
				39	0.003026		

Appendix A

Visual Basic Code

```
Attribute VB_Name = "Module1"

Option Base 1

'Lambda and theta are the parameters of generalized Poisson

' distribution

'DIST hold the index for the appropriate claim size distribution

' Parameter(3) holds upto three parameters of claim size distribution

' Span is the span of the claim sizes distribution

Public Lambda As Double, Theta As Double

Public DistributionList(4) As String

Public DIST As Integer, Span As Integer, Nmax As Long

Public Parameter(3) As Double

Public Vector() As Double, Compound() As Double


'This sub displays appropriate forms then get the user input


Sub GetData()

 GeneralizedPoisson.Show

 DistributionList(1) = "Gamma Distribution"
```

```
 DistributionList(2) = "Log Normal Distribution"
 DistributionList(3) = "Weibull Distribution"
 DistributionList(4) = "Burr Distribution"
 Claimsizes.Show
 Select Case DIST
  Case 1
   Gamma.Show
  Case 2
   Lognormal.Show
  Case 3
   Weibull.Show
  Case 4
   Burr.Show
  End Select
End Sub
Private Sub Workbook_Open()
 On Error Resume Next
 Worksheets("Report").Add
 Worksheets("Report").Active
End Sub
```

```
Sub Main()

  Call GetData

  Call Discretization(DIST)

  Call Recur(Nmax, Compound)

End Sub


'This sub discretize one of the chosen continous distributions


Sub Discretization(DIST As Integer)

 Dim i As Long, Alpha As Double, Beta As Double, Xmax As Double

 Dim x1 As Double, x2 As Double, Nmax1 As Long

 Alpha = Parameter(1)

 Beta = Parameter(2)

 Xmax = INVDistribution(DIST, (0.9999))

 Nmax = Xmax / Span + 1

 If Nmax > 60000 Then Nmax = 60000

 ReDim Vector(Nmax, 4)

 x2 = Span

 Vector(1, 1) = 0

 Vector(1, 2) = x2
```

```
 Vector(1, 3) = Distribution(DIST, x2)
 Vector(1, 4) = Vector(1, 3)
 For i = 2 To Nmax
  x1 = (i - 1) * Span
  x2 = i * Span
  Vector(i, 1) = x1
  Vector(i, 2) = x2
  Vector(i, 3) = Distribution(DIST, x2) - Distribution(DIST, x1)
  Vector(i, 4) = Vector(i - 1, 4) + Vector(i, 3)
 Next
 Range("a1:I2000").ClearContents
 With Range("A1")
  Range(.Offset(0, 0), .Offset(Nmax - 1, 3)).Name = "Vector"
 End With
 Range("Vector").Value = Vector


End Sub
Function INVDistribution(DIST As Integer, Prob As Double) As Double
Select Case DIST
 Case 1
```

```
 INVDistribution = Application.WorksheetFunction.GammaInv(Prob, (Parameter(1)),
 (Parameter(2)))

Case 2

 INVDistribution = Application.WorksheetFunction.LogInv(Prob, (Parameter(1)),
 (Parameter(2)))

Case 3

  INVDistribution = (-Log(1 - Prob) / Parameter(1)) ^ (Parameter(2))

Case 4

 INVDistribution = (((1 - Prob) ^ (-1 / Parameter(1)) - 1) *
 Parameter(2)) ^ (1 / Parameter(3))

End Select

End Function

Function Distribution(DIST As Integer, x As Double)

Select Case DIST

Case 1

 Distribution = Application.WorksheetFunction.GammaDist(x, Parameter(1),
 Parameter(2), True)

Case 2

 Distribution = Application.WorksheetFunction.LogNormDist(x, Parameter(1),
 Parameter(2))

Case 3

 Distribution = 1 - Exp(-Parameter(1) * x ^ Parameter(2))
```

```
 Case 4
  Distribution = 1 - (Parameter(2) / (Parameter(2) + x ^ Parameter(3))) ^
  Parameter(1)
End Select
End Function
Sub Recur(Nup As Long, Compound As Variant)
 Dim i As Integer, j As Integer, Temp As Double, y As Integer
 Dim Nup1 As Long
 Dim ScaleFactor As Double
 ScaleFactor = 10 ^ 300
 Nup1 = (300 * Log(10) - Lambda) / Theta
 If Nup > Nup1 Then
  Nup = Nup1
 Else
  ScaleFactor = 1
 End If
 ReDim Compound(Nup + 1, 2)
 ReDim CGPD(Nup + 1, Nup + 1)
 For j = 1 To Nup
   CGPD(1, j) = Exp(-Lambda - (j - 1) * Theta) * ScaleFactor
 Next
```

```
 For i = 1 To Nup
  For j = 0 To Nup - i
    Temp = 0
    For y = 1 To i
     Temp = Temp + (Theta + (Lambda + j * Theta) * y / i) * CGPD(i + 1 - y,
     j + 2) * Vector(y, 3)
    Next
    CGPD(i + 1, j + 1) = Temp * (Lambda + j * Theta) / (Lambda +
    (j + 1) * Theta)
   Next
 Next
For i = 1 To Nup + 1
 Compound(i, 1) = i - 1
 Compound(i, 2) = CGPD(i, 1) / ScaleFactor
Next
With Range("G1")
 Range(.Offset(0, 0), .Offset(Nup, 1)).Name = "Compound"
End With
Range("Compound").Value = Compound
End Sub
```

Bibliography

[1] Ambagaspitiya, R. S.(1995). A family of discrete distributions. *Insurance: Mathematics and Statistics*, 16, 107-127.

[2] Ambagaspitiya, R. S. and Balakrishan, N. (1994). On the compound generalized poisson distributions. *ASTIN Bulletin*, Volume 24, No.2, 255-263.

[3] Bowers, N., Gerber, H., Hickman, J., Jones, D., and Nesbitt, C. (1997). *Actuarial Mathematics*, 2nd edition, Society of Actuaries.

[4] Burr, I. W. (1942). Cumulative frequency functions. *Annals of Mathematical Statistics*, 13, 215-232.

[5] Consul, P. C. (1990). A modal for distributions of injuries in auto-accidents. *Itteilungen der Schweiz Vereinigung der Versicherungsmathematiker* **Helf 1**, 161-168.

[6] Consul, P. C. and Gupta, H. S. (1980). The generalized negative binomial distribution and its characterization by zero regression. *SIAM Journal of Applied Mathematics* 39, 231-237.

[7] Consul, P. C.and Jain, G. C. (1973). A generalization of the Poisson distribution. *Technometrics* 15, 791-799.

[8] Consul, P. C. and Shoukri, M. M. (1985). The generalized Poisson distribution when the sample mean is larger than the sample variance. *Communication in Statistics - Simulation and Computation* 14, 1533-1547.

[9] Douglas, J. B. (1971). Stirling numbers in discrete distributions. *Statistical Ecology 1: Spatial patterns and Statistical Distributions*, G. P. Patil, E. C. pielou, and W. E. Waters (editors), 69-98. University Park: Pennsylvania state University Press.

[10] Douglas, J. B. (1980). *Analysis with Standard Contagious Distributions.* Burtonsville, MD: International Co-operative Publishing House.

[11] Feller, W. (1943). On a general class of "contagious" distributions. *Annals of mathematical Statistics*, 14, 389-400.

[12] Galton, F. (1879). The geometri mean in vital and social statistics. *Proceedings of the Royal Society of London*, 29, 365-367.

[13] Gilbart, R. (1930). UNe loi des répartitions économiques; L'effet proportionelle, Bulletin de statistique Géneral, France, 19 469ff.

[14] Hagstroem, K. G. (1960) Remarks on Pareto distributions, *Skandinavisk Aktuarietidskrift*, 43, 59-71.

[15] Harris, C. M. (1968). The Pareto distribution as a queue service discipline, *Operations Research*, 16, 2155-2164.

[16] Hayakawa, M. (1951). The application of Pareto's law of income to Japanese data, *Econometrica*, 19, 174-183.

[17] Heckman, P. and Meyers, G. (1983). The Calculation of Aggregate Loss Distributions from Claim Severity and Claim count Distributions. *Proceedings of the Casualty Actuarial Society*, 70, 22-61.

[18] Jain, G. C. and P. C. Gupta (1971). A generalized negative binomial distribution. *SIAM Journal of Applied Mathematics*, 21, 501-503

[19] Johnson, N. L., Kotz, S. and Kemp A. W. (1992). *Univariate Distributions*, Second Edition, John Wiley & Sons, Inc.

[20] Johnson, N. L., Kotz, S. and Balakrishan, N. (1994). *Continuous Univariate Distributions*, Volume 1, Second Edition, John Wiley & Sons, Inc.

[21] Kapteyn, J. C. (1930). *Skew Frequency Curves in Biology and Statistics*, Groningen: Astronomical Laboratory, Noordhhoft.

[22] Kass, R. and Goovaerts, M. J. (1985). Application of the problem of moments to various insurance problems in non-life. In *Proceedings of the NATO Advanced Study on Insurance and Risk Theory*, edited by Goovaerts, M., de Vylder, F. and Haezendonck, J.. NATO ASI Series, 79-118.

[23] Maguire, B. A., Pearson, S. E. and Wynn, H. A. (1952). The time intervals between industrial accidents, *Biometrika*, 39, 168-180.

[24] Mandelbrot, B. (1960). The Pareto-Levy law and the distribution of income, *International Economics Review*, 1, 79-106.

[25] Mandelbrot, B. (1963). The variation of certain speculative prices, *Journal of Business, University of Chicago*, 36, 394-419.

[26] Mandelbrot, B. (1967). The variation of some other speculative prices, *Journal of Business, University of Chicago*, 40, 393-413.

[27] McAlister, D. (1979). The law of the geometri mean, *Proceedings of the Royal Society of of London*, 29, 367-375.

[28] Medgyessy, P. (1977). *Decomposition of superpositions of Density Functions and Discrete Distributions.* Budapest: Akademia Kiado; Bristol: Adam Hilger.

[29] Neyman, J.(1939). On a new class of "contagious" distributions applicable in entomology and bacteriology. *Annals of Mathematical Statistics*, 10, 35-57.

[30] Ord, J. K. (1975). Statistical models for personal income distributions, In *Statistical Distributions in Scientific work*, 2, G. P. Patial, S. Kotz, and J. K. Ord (editors), 151-158, Dordrecht: Reidel.

[31] Panjer, H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin.*

[32] Panjer, H.H., and Willmot, G.E. (1992). *Insurance Risk Models.* Society of Actuaries.

[33] Pigou, A. C. (1932). *The Economics of Welfare*, London: Macmillan.

[34] Seal, H. L. (1978a). From aggregate claims distribution to probbability of ruin. *ASTIN Bulletin*, 10, 47-53.

[35] Shirras, G. F. (1935). The Pareto law and the distribution of income, *economic Journal*, 45, 663-681.

[36] Steindl, J. (1965). *Random Processes and the Growth of firms*, New York: Hafner.

[37] Stuart, A and Ord, J. K. (1994). *Kendall's Advanced Theory of Statistics*, Volume 1, 6th edition, John Wiely.

[38] Sundt, B. and W. S. Jewell (1981) Further results on recursive evaluation of compound distributions. *ASTIN Bulletin*, 12, 27-39.

[39] Uven, M. J. van (1971a). Logarithmic frequency distributions, *Proceedings of the Royal Academy of Science, Amsterdam*, 19, 670-694.

[40] Weibull, W. (1939a). A ststistical theory of the strength of material, Report No. 151, Ingeniörs Vetenskaps Akademines Handligar, Stockholm.

[41] Weibull, W. (1939b). The phenomenon of rupture in solids, Report No. 153, Ingeniörs Vetenskaps Akademines Handligar, Stockholm.