PRIZE-WINNING PAPERS

SUNDT AND JEWELL'S FAMILY OF DISCRETE DISTRIBUTIONS

BY GORDON WILLMOT

University of Waterloo

ABSTRACT

A class of claim frequency distributions discussed by SUNDT and JEWELL (1981) is completely enumerated. Computational techniques for the associated compound total claims distribution in the presence of policy modifications are then derived.

KEYWORDS

Panjer's recursion; deductibles; maximums; extended truncated negative binomial.

1. INTRODUCTION

The total claims payable on a portfolio of business is often modelled as a random sum, or a compound distribution, in order to account for randomness in both frequency and severity of claims. Computation of the distribution of total claims often causes difficulty, but for certain parametric claim-frequency distributions the probability density function (pdf) may be obtained numerically as the solution to an integral equation. See PANJER (1981), SUNDT and JEWELL (1981), or STROTER (1985), for details.

The total claims distribution may be complicated by the imposition of certain policy modifications, such as deductibles and maximums This has the effect of creating a more complicated claim-severity distribution, for which the usual integral equation does not hold.

It is the aim of this paper to study a family of number of claims distributions which was introduced by SUNDT and JEWELL (1981), and the associated total-claims distribution. All members of the family are enumerated, and then an invariance property of a larger family is derived which leads to distributional and computational simplifications in the presence of certain types of reinsurance.

A more general procedure for Sundt and Jewell's family is then outlined which allows one to entertain models with a maximum benefit per claim. This may be used in conjunction with the previously mentioned invariance property to allow for the simultaneous treatment of deductibles and maximums.

In the final section, the extremely good fit to automobile claim frequency data of one member is demonstrated. The simplicity and flexibility of this distribution suggests that it be considered as a claim-frequency model.

2. BACKGROUND AND NOTATION

Let the number of claims N have probability distribution $\{p_n = \Pr(N = n), n = 0, 1, 2, \ldots\}$ and probability generating function (pgf) $P(z) = E(z^N)$. The claim sizes are denoted by a sequence $\{X_1, X_2, X_3, \ldots\}$ of non-negative independent and identically distributed random variables with Laplace–Stieltjes transform $L_X(s) = E(e^{-sX})$ where X is a generic claim size random variable. The total claims Y is defined by $Y = X_1 + X_2 + \ldots + X_N$ if $N \geqslant 1$ and 0 otherwise. The associated Laplace–Stieltjes transform is $L_Y(s) = E(e^{-sY})$, and it is well known that $L_Y(s) = P\{L_X(s)\}$.

SUNDT and JEWELL (1981) considered the family of number of claims distributions which satisfy the recursive relationship

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}; \qquad n = 2, 3, 4, \ldots \tag{2.1}$$

and show that if the claim sizes are absolutely continuous with pdf $f(x)$ for $x > 0$, then the pdf $g(x)$ of the total claims satisfies the integral equation

$$g(x) = p_1 f(x) + \int_0^x \left(a + b\frac{y}{x}\right) f(y) g(x-y)\, dy, \quad x > 0, \tag{2.2}$$

which may be solved numerically for $g(x)$. Thus, the relation (2.2), when combined with the fact that $\Pr(Y = 0) = p_0$, specifies the distribution of Y in this case.

3. MEMBERS OF THE CLASS

As discussed in SUNDT and JEWELL (1981), the family (2.1) includes the well-known Poisson, negative binomial, binomial, and logarithmic series distributions. Another member is the so-called extended truncated negative binomial (ETNB) distribution introduced by ENGEN (1974) with probability function

$$q_n = \frac{-\alpha\Gamma(n+\alpha)}{n!\Gamma(1+\alpha)} \frac{p^n}{1-(1-p)^{-\alpha}}; \qquad n = 1, 2, 3, \ldots, \tag{3.1}$$

and pgf

$$Q(z) = \frac{1-(1-pz)^{-\alpha}}{1-(1-p)^{-\alpha}}, \tag{3.2}$$

where $-1 < \alpha < 0$ and $0 < p < 1$. It will be demonstrated in Theorem 1 below that (3.1) is also a valid probability distribution if $p = 1$. Furthermore, if $\alpha > 0$ then the resulting distribution is a negative binomial truncated at 0, and so the possible parameter values are $\{-1 < \alpha < 0, 0 < p \leqslant 1\}$ and $\{0 < \alpha, 0 < p < 1\}$. The case with $\{-1 < \alpha < 0, 0 < p \leqslant 1\}$ shall henceforth be referred to as the ETNB distribution. As $\alpha \to 0$ it is easily shown that the logarithmic series distribution results. This distribution often provides a good fit to data and is discussed in more detail in Section 6.

From (2.1) and (3.1), one finds that $a = p$ and $b = p(\alpha - 1)$. SUNDT and JEWELL (1981) appear to have overlooked the possibility of a distribution with $a = 1$.

All members of the class (2.1) are identified in the following theorem.

THEOREM 1. *The only nondegenerate distributions which satisfy (2.1) are those with pgf of the form*

$$P(z) = \rho + (1 - \rho)Q(z) \tag{3.3}$$

with $-Q(0)[1 - Q(0)]^{-1} \leqslant \rho < 1$ *and* $Q(z)$ *the pgf of a Poisson, negative binomial, binomial, logarithmic series, or ETNB distribution.*

PROOF. SUNDT and JEWELL (1981) demonstrate that the only nondegenerate distributions which satisfy (2.1) are those with pgf of the form (3.3) where $Q(z)$ is the pgf of a Poisson, negative binomial, or binomial distribution, plus those which satisfy (2.1) with $a > 0$, $a + b \leqslant 0$, and $2a + b \geqslant 0$. If $2a + b = 0$ then, from (2.1), $p_2 = 0$, and the two-point binomial distribution (degenerate if $p_0 = 0$) results. If $a > 1$, then there must be a positive integer $N \geqslant 2$ such that $a + bN^{-1} = 0$ so that $p_N = 0$; otherwise there would exist m such that $p_{n+1} > p_n$ for all $n > m$, and the probabilities would diverge. Thus $b = -Na$, and from (2.1) one finds that $p_2 = a(1 - N/2)p_1$ is negative unless $N = 2$. But this means $2a + b = 0$ as before, and so $a \leqslant 1$ otherwise. From (2.1), one finds that

$$p_n = \frac{p_1}{n!} \frac{\Gamma(n + 1 + ba^{-1})}{\Gamma(2 + ba^{-1})} a^{n-1}; \qquad n = 1, 2, 3, \ldots \tag{3.4}$$

Also,

$$\frac{p_n}{p_{n-1}} = a\left(1 + \frac{ba^{-1}}{n}\right); \qquad n = 2, 3, 4, \ldots. \tag{3.5}$$

Then (letting $f_{n+1} = p_n$), it follows by Raabe's test (MARSDEN, 1974, p 60) that the series $\Sigma f_n = \Sigma p_n$ converges if $0 < a \leqslant 1$ and $a + b < 0$, since, in this case, from (3.5),

$$\frac{f_{n+1}}{f_n} \leqslant 1 - \frac{A}{n}$$

where $A = -ba^{-1} > 1$. Thus, let $p = a$ and $\alpha = 1 + ba^{-1}$. Thus $0 < p \leqslant 1$, and since $ba^{-1} < -1$ it follows that $\alpha < 0$. The condition $2a + b > 0$ implies that $\alpha > -1$. From (3.4), it follows that

$$p_n = \frac{p_1}{n!} \frac{\Gamma(n + \alpha)}{\Gamma(1 + \alpha)} p^{n-1}; \qquad n = 1, 2, 3, \ldots. \tag{3.6}$$

Now

$$1 = p_0 + \frac{p_1}{p}\sum_{n=1}^{\infty}\frac{\Gamma(n+\alpha)}{n!\Gamma(1+\alpha)}p^n$$

$$= p_0 + \frac{p_1}{\alpha p}\sum_{n=1}^{\infty}\binom{\alpha+n-1}{n}p^n$$

$$= p_0 + \frac{p_1}{\alpha p}\sum_{n=1}^{\infty}\binom{-\alpha}{n}(-p)^n$$

$$= p_0 + \frac{p_1}{\alpha p}\{(1-p)^{-\alpha}-1\}$$

Hence,

$$p_1 = \frac{-\alpha p(1-p_0)}{1-(1-p)^{-\alpha}}$$

and substitution in (3.6) yields

$$(3.7)\qquad p_n = (1-p_0)\left\{\frac{-\alpha\Gamma(n+\alpha)}{n!\Gamma(1+\alpha)}\,\frac{p^n}{1-(1-p)^{-\alpha}}\right\},\qquad n = 1,2,3, \ldots$$

Thus, (3.7) corresponds to (3.3) with $Q(z)$ the ETNB pgf (3.2), as may be seen by comparison with (3.1). Now suppose that $a+b=0$. Then, from (3.4),

$$(3.8)\qquad p_n = \frac{p_1 a^{n-1}}{n};\qquad n = 1,2,3, \ldots$$

Clearly, this series diverges if $a = 1$, and converges if $0 < a < 1$ since, for example, $p_n/p_{n-1} = a(1-1/n) \to a$ as $n \to \infty$, which implies convergence by the ratio test (MARSDEN, 1974, p. 47). From (3.8), one finds that for $0 < a < 1$,

$$1 = p_0 + \frac{p_1}{a}\sum_{n=1}^{\infty}\frac{a^n}{n}$$

$$= p_0 - \frac{p_1}{a}\log(1-a).$$

Thus, $p_1 = (1-p_0)\{a/(-\log(1-a))\}$ and (3.8) becomes

$$(3.9)\qquad p_n = (1-p_0)\left\{\frac{a^n}{-n\log(1-a)}\right\};\qquad n = 1,2,3,\ldots,$$

which corresponds to (3.3) with $Q(z)$ a logarithmic series pgf.

Summarizing, the line $2a+b=0$ corresponds to the two-point binomial distribution (degenerate if $p_0 = 0$). The region $\{0 < a \leqslant 1,\ a+b<0,\ 2a+b>0\}$ corresponds to the ETNB distribution (3.1) with $\{-1<\alpha<0, 0<p\leqslant 1\}$. The region $\{0<a<1,\ a+b=0\}$ corresponds to the logarithmic series distribution. There are no other possible regions.

If one sets $\rho = -Q(0)[1-Q(0)]^{-1}$ in (3.3), one obtains the truncated distribution, with the zero class missing.

Some of the other members have been used in insurance contexts. GOSSIAUX and LEMAIRE (1981) used the modified geometric in fitting automobile insurance data JEWELL and SUNDT (1981) suggested the use of the modified binomial in approximating the individual risk model. WANI and LO (1986) have considered the family (2.1) within the class of power series distributions, suggesting that the empirical ratio (2.1), obtained by replacing the probability p_n by the proportion of observations equal to n, may be used to discriminate between members of the family. By Theorem 1, it is clear that it may also be used to decide whether the family itself is appropriate.

4. DEDUCTIBLES AND REINSURANCE

There are many types of insurance agreements which give rise to models for the total claims as discussed earlier, but with a possible mass point f_0 at 0 and continuous density $f(x)$ for $x > 0$ for the amount actually payable on each claim. These include deductibles and excess-of-loss and certain catastrophe reinsurances. See PANJER and WILLMOT (1984) for details.

It is convenient to introduce the conditional pdf $f_c(x) = f(x)/(1 - f_0)$ with Laplace transform ($f_c(x) = 0$ if $x < 0$)

$$\tilde{f}_c(s) = \int_0^\infty e^{-sx} f_c(x)\, dx \tag{4.1}$$

Then the Laplace–Stieltjes transform of the single-claim amount distribution becomes

$$L_X(s) = f_0 + (1 - f_0)\tilde{f}_c(s), \tag{4.2}$$

and that of the total claims is thus

$$L_Y(s) = P\{f_0 + (1 - f_0)\tilde{f}_c(s)\}. \tag{4.3}$$

The recursive formula (2.2) is not applicable in this case owing to the presence of the mass point f_0 at 0. However, the distribution with transform (4.3) may be expressed in a fashion which permits the use of (2 2) with little additional difficulty when the number-of-claims distribution satisfies (2.1) In particular, note that (3.3) may be rewritten as

$$P(z) = P(0) + [1 - P(0)]K(z) \tag{4.4}$$

where $K(z) = \{Q(z) - Q(0)\}/\{1 - Q(0)\}$. Since the relation (2.1) begins at $n = 2$, $P(0)$ is a free parameter. Similarly, for all members of the class (2.1), one can find a parameter λ and a function $B(x)$ not depending on λ such that

$$K(z;\ \lambda) = \frac{B[\lambda(z-1)] - B(-\lambda)}{1 - B(-\lambda)}, \tag{4.5}$$

where the explicit dependency on λ is noted on the left-hand side of (4 5).

For the Poisson case, $B(x) = e^x$ and $\lambda > 0$ For the negative binomial, $B(x) = (1 - x)^{-\alpha}$ with $\alpha > 0$ and $\lambda > 0$. The ETNB pgf (3 2) is of the same form as the negative binomial but with $-1 < \alpha < 0$ and $\lambda = p(1-p)^{-1} > 0$ (unless $p = 1$, in which case one may choose $B(x) = 1 - (-x)^{-\alpha}$ where $-1 < \alpha < 0$ and λ is arbitrary since (4.5) does not involve λ). The binomial has $B(x) = (1+x)^n$ where $0 < \lambda < 1$ and n is a positive integer. One has $B(x) = 1 + \log(1-x)$ and $\lambda > 0$ for the logarithmic series distribution.

Hence, the pgf (3.3) depends on (at least) two parameters π and λ, and thus it may be expressed as

$$P(z;\ \lambda,\ \pi) = \pi + (1-\pi)K(z;\ \lambda) \tag{4.6}$$

where $K(z;\ \lambda)$ is given by (4.5) and $0 \leqslant \pi < 1$. The main result of this section is the following, which generalizes that of PANJER and WILLMOT (1984) and includes all members of the family (2.1) as special cases.

THEOREM 2. *For the pgf (4.6),*

$$P\{f_0 + (1-f_0)z;\ \lambda, \pi\} = P\{z;\ \lambda(1-f_0),\ P(f_0;\ \lambda, \pi)\}. \tag{4.7}$$

PROOF. From (4.5),

$$\begin{aligned}
K\{f_0 + (1-f_0)z;\ \lambda\} &= \frac{B\{\lambda(1-f_0)(z-1)\} - B(-\lambda)}{1 - B(-\lambda)} \\
&= \frac{B[-\lambda(1-f_0)] - B(-\lambda)}{1 - B(-\lambda)} + \frac{1 - B[-\lambda(1-f_0)]}{1 - B(-\lambda)} \\
&\quad \times \left\{\frac{B[\lambda(1-f_0)(z-1)] - B[-\lambda(1-f_0)]}{1 - B[-\lambda(1-f_0)]}\right\} \\
&= K(f_0;\ \lambda) + [1 - K(f_0;\ \lambda)]K\{z;\ \lambda(1-f_0)\}.
\end{aligned}$$

Thus, from (4.6),

$$\begin{aligned}
&P\{f_0 + (1-f_0)z;\ \lambda, \pi\} \\
&\qquad = \pi + (1-\pi)K\{f_0 + (1-f_0)z;\ \lambda\} \\
&\qquad = \pi + (1-\pi)\{K(f_0;\ \lambda) + [1 - K(f_0;\ \lambda)]K[z;\ \lambda(1-f_0)]\} \\
&\qquad = P(f_0;\ \lambda, \pi) + [1 - P(f_0;\ \lambda, \pi)]K\{z;\ \lambda(1-f_0)\}.
\end{aligned}$$

One consequence of Theorem 2 is the fact that the total number of claims which are of a particular amount (or range of amounts) is from the same family of distributions as the total number of claims This follows by letting f_0 be the probability that a particular claim is not of the amount of interest, and then (4.7) is the pgf of the number of claims of the amount of interest.

Suppose an insurer only observed claims of a certain size (for example, a deductible may cause an individual not to submit claims which are clearly below the deductible level), and it was decided to fit a number of claims distribution to these observed claims only. If this model does not satisfy (4.7), then the implied total number-of-claims distribution is not from the same family. This implies that a different model would have been selected had the deductible been omitted (or had it been different). This is likely to be undesirable, and so (4.7) may be considered to be an important property.

From a computational standpoint, the theorem is also convenient, since the recursion (2 2) may be applied (if, in addition, (2.1) is satisfied) after first applying the theorem. Simply replace $f(x)$ by $f_c(x)$ and use the same number-of-claims distribution as before, but with $p_0 = \pi$ replaced by $P(f_0; \lambda, \pi)$ and the parameter λ replaced by $\lambda(1 - f_0)$, as may be seen from (4.3) and substitution of $\tilde{f}_c(s)$ for z in (4.7). This is convenient in that the recursive formula need not be modified. PANJER and WILLMOT's (1984) result is recovered by choosing $\pi = B(-\lambda)$, but is only applicable if $B[\lambda(z-1)]$ is itself a pgf

It is clear that the distributions which satisfy (4.7) are invariant under random-sum reinsurance agreements. WANI and LO (1983) describe distributions which satisfy (4.7) in a biological context, referring to them as invariant abundant distributions. They show that the only power series distributions which satisfy (4 7) are all the members of the class (2.1). Theorem 2 generalizes this family as well, since it depends on the representation (4.6) only, and is not restricted to power series distributions and thus the family defined by (2.1).

5 MAXIMUM BENEFITS PAYABLE

In the previous section, methods for dealing with claims of size 0 were presented. A similar problem which is more difficult to deal with mathematically involves the imposition of a maximum benefit payable. This has the effect of creating a single-claim-amount distribution with a single mass point and a density portion.

Hence, it is now assumed that the single-claim-amount distribution has a mass point f_m at the value m, and conditional pdf $f_c(x)$, given that m is not payable (again it is assumed that $f_c(x) = 0$ if $x < 0$) This generalizes the assumption of the previous section, where it is assumed that $m = 0$.

If the underlying claim size distribution has pdf $f(x)$, then the imposition of a maximum benefit of m implies that

$$f_m = \int_m^\infty f(x)\,dx \tag{5.1}$$

and

$$f_c(x) = \begin{cases} f(x)/(1 - f_m), & 0 < x < m \\ 0, & x \geqslant m. \end{cases} \tag{5.2}$$

Thus, a situation involving a maximum may be treated as a special case of what follows.

Similarly, a deductible may be handled by letting $m = 0$ in the following, thus providing an alternative derivational approach to that of the previous section. However, in the situation involving both a deductible and a maximum, it is much more convenient to "remove the mass point at 0" using the results of the previous section and then use the following for the maximum only, rather than allowing for two mass points (one of which is at 0) and a continuous portion. The latter approach leads to unnecessarily complicated algebra.

The total claims Y has mass points at integral multiples of m, i.e.

$$\Pr(Y = mn) = p_n f_m^n; \qquad n = 0, 1, 2, \ldots, \tag{5.3}$$

and $P(f_m)$ is the total of the discrete portion (if $m = 0$, there is a single mass point $P(f_0)$ at 0). The pdf portion may be obtained by conditioning on the number of claims, and then on the number of nonzero claims. Therefore let

$$g_n(x) = \sum_{k=1}^{n} \binom{n}{k} f_m^{n-k}(1 - f_m)^k f_c^{*k}[x - m(n-k)] \tag{5.4}$$

where $f_c^{*k}(x)$ is the k-fold convolution of $f_c(x)$ with itself. Thus, $g_n(x)$ is the conditional pdf of Y given that n claims occurred, and is obtained by conditioning on how many are for amount m. Note that if $f_m = 0$, $g_n(x) = f_c^{*n}(x)$, as it must. Then the pdf portion of the distribution of Y is

$$g(x) = \sum_{n=1}^{\infty} p_n g_n(x). \tag{5.5}$$

Clearly, (5.5) is not well suited for computational purposes, and it is of interest to derive a computational formula for $g(x)$. For any number-of-claims distribution with pgf satisfying (4.7), the total claims of size m and of those not equal to m may be obtained easily as discussed in that section, but only in the Poisson case are they independent (cf. KARLIN and TAYLOR, 1981, pp. 433–6). Thus, only in the Poisson case may a standard convolution approach be used to compute the total claims distribution. Hence, a more general approach is needed for the class (2.1), and a generalization of (2.2) is now given.

THEOREM 3. *If the claim frequency distribution satisfies (2.1), then the density (5.5) satisfies the integral equation*

$$\begin{aligned} g(x) = {} & p_1(1 - f_m)f_c(x) + (1 - f_m) \sum_{n=1}^{[x/m]} p_n \left[a + b\left(1 - \frac{mn}{x}\right)\right] f_m^n f_c(x - mn) \\ & + \left(a + b\,\frac{m}{x}\right) f_m g(x - m) + (1 - f_m) \int_0^x \left(a + b\,\frac{y}{x}\right) f_c(y) g(x - y)\, dy, \end{aligned} \tag{5.6}$$

where $[x]$ *denotes the greatest-integer function.*

PROOF. Suppose that, for fixed n, $X_1, X_2, \ldots, X_n$ are iid single claim amount random variables. Then by symmetry, one has that

$$a + \frac{b}{n} = E\left\{a + b\,\frac{X_1}{x} \,\middle|\, \sum_{i=1}^{n} X_i = x\right\}. \tag{5.7}$$

If $x \neq km$, then the right-hand side of (5.7) may be obtained by considering the case when $X_1 = m$, all but X_1 are equal to m, and when X_1 and others are equal to m (note that not all may be equal to m simultaneously). This yields

$$\begin{aligned}\left(a+\frac{b}{n}\right)g_n(x) &= \left(a+b\,\frac{m}{x}\right)f_m g_{n-1}(x-m)\\ &+\left\{a+\frac{b}{x}\,[x-m(n-1)]\right\}(1-f_m)f_m^{n-1}f_c\,[x-m(n-1)]\\ &+(1-f_m)\int_0^x\left(a+b\,\frac{y}{x}\right)f_c(y)g_{n-1}(x-y)\,dy\end{aligned}\tag{5.8}$$

Thus, from (5.5), (2.1), and (5.8), one finds that

$$\begin{aligned}g(x) &= \sum_{n=1}^{\infty} p_n g_n(x) = p_1 g_1(x) + \sum_{n=2}^{\infty} p_{n-1}\left(a+\frac{b}{n}\right)g_n(x)\\ &= p_1(1-f_m)f_c(x) + \sum_{n=2}^{\infty} p_{n-1}\left\{a+b\,\frac{m}{x}\right\}f_m g_{n-1}(x-m)\\ &+\sum_{n=2}^{\infty} p_{n-1}\left\{a+\frac{b}{x}\,[x-m(n-1)]\right\}(1-f_m)f_m^{n-1}f_c\,[x-m(n-1)]\\ &+\sum_{n=2}^{\infty} p_{n-1}(1-f_m)\int_0^x\left(a+b\,\frac{y}{x}\right)f_c(y)g_{n-1}(x-y)\,dy\\ &= p_1(1-f_m)f_c(x) + \left(a+b\,\frac{m}{x}\right)f_m\sum_{n=2}^{\infty} p_{n-1}g_{n-1}(x-m)\\ &+(1-f_m)\sum_{n=2}^{\infty} p_{n-1}\left\{a+b\left[1-\frac{m(n-1)}{x}\right]\right\}f_m^{n-1}f_c\,[x-m(n-1)]\\ &+(1-f_m)\int_0^x\left(a+b\,\frac{y}{x}\right)f_c(y)\left\{\sum_{n=2}^{\infty} p_{n-1}g_{n-1}(x-y)\right\}dy\\ &= p_1(1-f_m)f_c(x) + \left(a+b\,\frac{m}{x}\right)f_m g(x-m)\\ &+(1-f_m)\sum_{n=1}^{\infty} p_n f_m^n\left\{a+b\left[1-\frac{mn}{x}\right]\right\}f_c(x-mn)\\ &+(1-f_m)\int_0^x\left(a+b\,\frac{y}{x}\right)f_c(y)g(x-y)\,dy\end{aligned}$$

which is (5 6) since $f_c(x)=0$ if $x<0$.

The theorem may also be proved analytically along the lines of WILLMOT and PANJER (1987) using Laplace transforms, although the algebraic details are more cumbersome.

Some comments are in order at this point. First, if $f_m = 0$ then (5.6) reduces to (2.2). Also, if $m = 0$ then (5.6) reduces to

$$g(x) = \frac{1 - f_0}{1 - af_0}\left\{[p_1 + (a+b)(P(f_0) - p_0)]f_c(x)\right.$$

$$\left. + \int_0^x \left(a + b\,\frac{y}{x}\right) f_c(y) g(x-y)\, dy\right\}, \tag{5.9}$$

in agreement with (1.21) of WILLMOT (1986) Furthermore, if (5 2) holds, then $f_c(x - mn)$ is nonzero for but one or two values of n and so the complicated sum in (5.6) becomes simpler. Finally, it seems clear that this approach is tractable, yet cumbersome algebraically, if there are several mass points. Thus, for example, a situation involving both a deductible and a maximum leads to mass points at 0 and m, and it is simpler to "remove" the mass point at 0 using (4.7), and then use Theorem 3 to deal with m, rather than generalize Theorem 3 to the case involving two mass points.

6. THE MODIFIED ETNB DISTRIBUTION

Consider the distribution with pgf

$$P(z) = \pi + (1 - \pi)Q(z) \tag{6.1}$$

where $Q(z)$ is defined by (3.2) with parameter space $\{-1 < \alpha < 0, 0 < p \leqslant 1\} \cup \{0 < \alpha, 0 < p < 1\}$, and where $0 \leqslant \pi < 1$ The resultant probabilities are given by

$$p_0 = \pi \tag{6.2}$$

and

$$p_n = (1 - \pi)q_n; \qquad n = 1, 2, 3, \ldots, \tag{6.3}$$

where q_n is given by (3.1). This distribution is quite flexible owing to the extended range of the parameter space, and includes the modified geometric of GOSSIAUX and LEMAIRE (1981) as the special case $\alpha = 1$, and the logarithmic series with zeros as the limiting case $\alpha \to 0$

The integral equation (2.2) may be augmented with a convenient asymptotic formula for the tail of the total-claims distribution Straightforward application of Stirling's formula (FELLER, 1968) to (3 1) yields

$$p_n \sim \frac{\alpha(1 - \pi)n^{\alpha - 1}}{\Gamma(1 + \alpha)[(1 - p)^{-\alpha} - 1]}\,p^n, \qquad n \to \infty. \tag{6.4}$$

This is an asymptotic formula of the form discussed by EMBRECHTS, *et al* (1985) Consequently, if $p < 1$ the tail of the associated compound distribution satisfies

$$Pr(Y > x) \sim \frac{\alpha(1 - \pi)x^{\alpha - 1}e^{-\kappa x}}{\kappa\Gamma(1 + \alpha)[(1 - p)^{-\alpha} - 1][-pL_X'(-\kappa)]^{\alpha}}, \qquad x \to \infty \tag{6.5}$$

where the single-claim-amount distribution is nonarithmetic and $\varkappa > 0$ satisfies $L_X(-\varkappa) = p^{-1}$. (This includes the situation discussed in the previous section involving the maximum benefit payable.) The formula (6.5) should be used in conjunction with (2.2) or (5.6), which are not as convenient for large values of x.

The distribution provides a very good fit to automobile insurance data. Suppose that the data consist of $\{F_k;\ k = 0, 1, 2, \dots\}$ where F_k represents the number of policies with k claims. The method of maximum likelihood suggests that the values of π, α, and p should be chosen so as to maximize

$$l(\pi, \alpha, p) = \sum_{k=0}^{\infty} F_k \log p_k. \tag{6.6}$$

For notational simplicity, let

$$N = \sum_{k=0}^{\infty} F_k \tag{6.7}$$

be the number of policies and

$$\bar{X} = N^{-1} \sum_{k=1}^{\infty} kF_k \tag{6.8}$$

be the average number of claims per policy. Then, setting the partials of (6.6) equal to 0 and solving for the parameters, one finds that the maximum-likelihood estimate of π is

$$\hat{\pi} = N^{-1} F_0 \tag{6.9}$$

and the maximum likelihood estimates of α and p satisfy the equations

$$\bar{X} = \frac{(1-\hat{\pi})\hat{\alpha}\hat{p}/(1-\hat{p})}{1-(1-\hat{p})^{\hat{\alpha}}} \tag{6.10}$$

and

$$-\frac{N\bar{X}}{\hat{p}}(1-\hat{p})\log(1-\hat{p}) = \hat{\alpha} \sum_{k=1}^{\infty} F_k \left\{ \sum_{m=0}^{k-1} (\hat{\alpha} + m)^{-1} \right\}. \tag{6.11}$$

The double sum on the right-hand side of (6.11) may also be written as $\sum_{m=0}^{\infty} (\hat{\alpha} + m)^{-1} \sum_{k=m+1}^{\infty} F_k$, a formula which may be more convenient for computational purposes.

The derivation of (6.10) and (6.11) is straightforward but tedious and follows that of ANSCOMBE (1950). Equation (6.10) equates the theoretical mean to the sample mean. The values of $\hat{\alpha}$ and $\hat{p}$ must be obtained numerically, but this causes little difficulty using a standard Newton–Raphson algorithm.

The distribution was fitted using this method to six automobile claim-frequency data sets given by GOSSIAUX and LEMAIRE (1981). The results are given in Table 1 where the fitted values are denoted by E_k. The Pearson goodness-of-fit statistic, associated degrees of freedom, significance level, and maximum-likelihood estimates of the parameters are also given. In some cases, grouping

TABLE 1

k	F_k	E_k	F_k	E_k
	Data Set 1		Data Set 2	
0	103,704	103,704 00	20,592	20,592 00
1	14,075	14,075 97	2,651	2,651 33
2	1,766	1,761 48	297	295 84
3	255	261 32	41	41 94
4	45	41 80	7	6 58
5	6	6 98	0	1 09
6	2	1 20	1	0 19
Total	119,853	119,852 75	23,589	23,588 96
Chi-squared	0 76		0 13	
DF	3		2	
Significance level	0 86		0 94	
$(\hat{\alpha}, \hat{p}, \hat{\pi})$	(0 285, 0 195, 0 865)		(0 104, 0 202, 0 873)	
	Data Set 3		Data Set 4	
0	370,412	370,412 00	3,719	3,719 00
1	46,545	46,546 57	232	232 09
2	3,935	3,929 04	38	37 26
3	317	323 73	7	8 51
4	28	26 35	3	2 23
5	3	2 13	1	0 63
Total	421,240	421,239 81	4,000	3,999 73
Chi-squared	0 46		0 52	
DF	2		1	
Significance level	0 80		0 47	
$(\hat{\alpha}, \hat{p}, \hat{\pi})$	(1 154, 0 078, 0 879)		(−0 119, 0 364, 0 930)	
	Data Set 5		Data Set 6	
0	7,840	7,840 00	96,978	96,978 00
1	1,317	1,320 31	9,240	9,241 89
2	239	225 19	704	696 12
3	42	54 14	43	53 49
4	14	14 91	9	4 15
5	4	4 42	0	0 35
6	4	1 37	—	—
7	1	0 44	—	—
Total	9,461	9,460 78	106,974	106,974 00
Chi-squared	8 03		6 64	
DF	3		1	
Significance level	0 05		0 01	
$(\hat{\alpha}, \hat{p}, \hat{\pi})$	(−0 103, 0 380, 0 829)		(0 886, 0 080, 0 907)	

was done to ensure that expected frequencies are sufficiently large (i e. greater than 1). The fit is quite good for the first four data sets, and reasonable for the last two. The apparently poor fit of data set 6 is deceptive in that the mismatch in cells 3 and 4 offset each other. If they were grouped, as GOSSIAUX and LEMAIRE (1981) did, the fit would be deemed adequate. It should be noted that $\hat{\alpha}$ was positive for data sets 1, 2, 3, and 6, suggesting that the truncated negative binomial with zeros was fit.

ACKNOWLEDGEMENTS

The author wishes to thank the referee for numerous suggestions which improved the presentation and readability of the paper. This work was supported by the National Sciences and Engineering Research Council of Canada.

REFERENCES

ANSCOMBE, F (1950) Sampling theory of the negative binomial and logarithmic series distributions *Biometrika* **37**, 358–382

EMBRECHTS, P, MAEJIMA, M, and TRUGELS, J (1985) Asymptotic behaviour of compound distributions *ASTIN Bulletin* **15**, 45–48

ENGEN, S (1974) On species frequency models *Biometrika* **61**, 263–270

FELLER, W (1968) *An Introduction to Probability Theory and its Applications*, Vol 1 (3rd edn) Wiley, New York

GOSSIAUX, A and LEMAIRE, J (1981) Methodes d'ajustement de distributions de Sinistres *Bulletin of the Association of Swiss Actuaries* **81**, 87–95

JEWELL, W and SUNDT, B (1981) Improved approximations for the distribution of a heterogeneous risk portfolio *Bulletin of the Association of Swiss Actuaries* **81**, 221–240

KARLIN, S and TAYLOR, H (1981) *A Second Course in Stochastic Processes* Academic Press, New York

MARSDEN, J (1974) *Elementary Classical Analysis* W H Freeman, San Fransisco

PANJER, H (1981) Recursive evaluation of a family of compound distributions *ASTIN Bulletin* **12**, 22–26

PANJER, H and WILLMOT, G (1984) Computational techniques in reinsurance models *Transactions of the 22nd International Congress of Actuaries*, Sydney, **4**, 111–120

STROTER, B (1985) The numerical evaluation of the aggregate claim density function via integral equations *Blatter der Deutschen Gesellschaft fur Versicherungsmathematik* **17**, 1–14

SUNDT, B and JEWELL, W (1981) Further results on recursive evaluation of compound distributions *ASTIN Bulletin* **12**, 27–39

WANI, J and LO, H (1983) A characterization of invariant power series abundance distributions *Canadian Journal of Statistics* **11**, 317–323

WANI J and LO, H (1986) Selecting a power-series distribution for goodness of fit *Canadian Journal of Statistics* **14**, 347–353

WILLMOT, G (1986) Mixed compound Poisson distributions *ASTIN Bulletin* **16S**, S59–S79

WILLMOT, G and PANJER, H (1987) Difference equation approaches in evaluation of compound distributions *Insurance Mathematics and Economics* **6**, 43–56

GORDON E. WILLMOT

University of Waterloo, Faculty of Mathematics, Department of Statistics and Actuarial Science, Waterloo, Ontario N2L 3G1, Canada.