ON AN INTEGRAL EQUATION FOR DISCOUNTED COMPOUND – ANNUITY DISTRIBUTIONS

BY COLIN M. RAMSAY

Actuarial Science, University of Nebraska, Lincoln NE, USA 68588-0307, (402) 472-5823

ABSTRACT

We consider a risk generating claims for a period of N consecutive years (after which it expires), N being an integer valued random variable. Let X_k denote the total claims generated in the k^{th} year, $k \geq 1$. The X_k's are assumed to be independent and identically distributed random variables, and are paid at the end of the year. The aggregate discounted claims generated by the risk until it expires is defined as $S_N(v) = \Sigma_{k=1}^{N} v^k X_k$, where v is the discount factor. An integral equation similar to that given by PANJER (1981) is developed for the *pdf* of $S_N(v)$. This is accomplished by assuming that N belongs to a new class of discrete distributions called annuity distributions. The probabilities in annuity distributions satisfy the following recursion:

$$p_n = p_{n-1}\left(a + \frac{b}{a_n}\right), \qquad \text{for} \qquad n = 1, 2, \ldots,$$

where a_n is the present value of an n-year immediate annuity.

KEYWORDS

Annuity distributions; integral equation; aggregate discounted claims.

1. INTRODUCTION

A major problem in mathematical risk theory is the evaluation of the distribution of the aggregate claims occuring in a fixed time period. This is because the aggregate claims is usually the sum of a random number of claims. If Y_k is the size of the k^{th} claim and N is the number of claims in this time period, then the aggregate claims S is given by

$$S = \sum_{k=1}^{N} Y_k. \tag{1}$$

The Y_k's are usually assumed to be independent and identically distributed (*iid*) with common cummulative distribution function (*cdf*) $F(y)$. If the n-fold convolution of $F(y)$ with itself is given by

$$F_n(y) = \int_0^y F_{n-1}(y-z)\,dF(z), \qquad n = 1, 2, \ldots,$$

with $F_0(y) = 1$, for $y \geq 0$, and the non-defective claim number distribution is

$$p_n = \Pr[N = n],$$

for $n = 0, 1, \ldots$, then the *cdf* of S is

$$G(y) = \sum_{n=0}^{\infty} p_n F_n(y). \tag{2}$$

Unfortunately, explicit expressions for $F_n(y)$ are usually not available, so the equation (2) is generally not very useful. Approximations for $G(y)$ are thus needed.

In order to facilitate the easy evaluation of $G(y)$ in equation (2), PANJER (1981), and SUNDT and JEWELL (1981) provided a family of claim number distributions which yielded an integral equation for the *pdf* of S when the Y_k's are absolutely continuous random variables. The random variable N must have probabilities satisfying the recursion

$$p_n = p_{n-1}\left(a + \frac{b}{n}\right) \tag{3}$$

where a and b are constants depending on the length of the time period. This family includes the geometric, Poisson, binomial, negative binomial, logarithmic series, and the so-called extended truncated negative binomial distribution. See WILLMOT (1988) for details. PANJER (1981) proved that if p_n satisfies equation (3), then $g(y)$, the *pdf* of S, satisfies the following integral equation for $y > 0$:

$$g(y) = p_1 f(y) + \int_0^y \left(a + \frac{bz}{y}\right) f(z)\, g(y-z)\, dz. \tag{4}$$

This integral equation can be solved numerically; see STRÖTER (1985).

Recall that S is defined as the aggregate claims over a fixed time period. If this time period T is large, i.e., extending over several years, then it many be prudent to include an interest discount factor to obtain the present value of these claims. Let T_k be the random time at which the claim Y_k occurs, and $N(T)$ be the number of claims over T years, T a positive integer. The aggregate discounted claims, denoted by $S_T^*(v)$, will be given by

$$S_T^*(v) = \sum_{k=1}^{N(T)} v^{T_k} Y_k \tag{5}$$

where $v = 1/(1+i)$ and i is the constant annual rate of interest. Comparing equations (1) and (5), it is clear that $S_T^*(v)$ is a more complicated random variable than S, and hence will have a more complicated *cdf*. $S_T^*(v)$ can be simplified by making the traditional actuarial assumption that claims are paid at the end of the year in which they occur. This means that equation (5) reduces to

$$S_T(v) = \sum_{k=1}^{T} v^k X_k \tag{6}$$

where X_k is the aggregate claims generated in year k. We assume that the number of claims occuring during each year is an *iid* sequence, implying that the X_k's are also *iid*.

The important observation to note here is that $S_T(v)$ is now the sum of T (a fixed number) of random variables X_k. Thus we have seen that the traditional model studied by PANJER and SUNDT and JEWELL can be adapted to include an interest factor. However an expression for the *pdf* of $S_T^*(v)$ will not be similar to equation (4) when the probabilities of $N(T)$ satisfy equation (3). We will see that by making T random, it is possible that $S_T(v)$ can be extended to yield a *pdf* which satisfies an integral equation similar to (4).

2. THE MAIN RESULTS

The inclusion of interest and/or inflation factors in risk theoretic models have appeared in the literature mainly in the context of the calculation of ruin probabilities; see, for example, WATERS (1983), BOOGAERTS and CRIJNS (1987), and GARRIDO (1988) and references therein. The limiting distributions of discounted processes have been studied by GERBER (1971), and BOOGAERT, HAEZENDONCK and DELBAEN (1988). However, there has been no work in the literature on integral equations similar to that of PANJER (1981) for aggregate discounted claims.

Consider a risk that can produce either no claim or it produces a sequence of *iid* positive claims that are paid at the end of the year in which they occured. Such risks are pertinent to health insurance, dental insurance, etc. The sequence of claims will run for N years, starting from year 1 until year N, after which no further claims are produced. N is an integer valued non-negative random variable. The total claims produced in the k^{th} year is $X_k > 0$, $k = 1, 2, \ldots$. If interest is at rate i annually, the aggregate discounted claims will be given by $S_N(v)$ where

$$S_N(v) = \sum_{k=1}^{N} v^k X_k \tag{7}$$

Notice the difference between equations (6) and (7), the constant T is now replaced by the random variable N. These equations clearly have different interpretations.

In order to develop an integral equation for the *pdf* of $S_N(v)$, we will introduce a new family of claim number distributions for N, called annuity distributions, with probabilities p_n satisfying the following difference equation:

$$(8) \qquad p_n = p_{n-1}\left(a + \frac{b}{a_n}\right), \qquad \text{for} \qquad n = 1, 2, \ldots,$$

where a_n is the present value of an n-year immediate annuity at interest rate i, i.e.,

$$(9) \qquad a_n = \frac{(1-v^n)}{i}.$$

As before, $p_n = \Pr[N = n]$.

Let $P(z)$ be the probability generating function of N, i.e.,

$$P(z) = \sum_{n=0}^{\infty} p_n z^n, \qquad \text{for} \qquad -1 \leq z \leq 1 .$$

It can easily be proven that

$$E[S_N(v)] = \frac{\mu(1-P(v))}{i}$$

and

$$\begin{aligned}\operatorname{Var}[S_N(v)] &= E[\operatorname{Var}[S_N(v) \mid N]] + \operatorname{Var}[E[S_N(v) \mid N]] \\ &= \frac{\sigma^2 v^2}{1-v^2}\left[1 - P(v^2) + \left(\frac{\mu}{i}\right)^2 [P(v^2)] - [P(v)]^2\right]\end{aligned}$$

where $\mu = E[X_k]$ and $\sigma^2 = \operatorname{Var}[X_k]$.

From equation (7) we condition on $\{N = n\}$ and define $S_n(v)$ as

$$S_n(v) = \sum_{k=1}^{n} v^k X_k, \qquad n = 1, 2, \ldots .$$

Note that, because the X_k's are *iid*, $S_n(v)$ has, for each non-negative integer m, the same distribution as

$$S_n(v) = \sum_{k=1}^{n} v^k X_{m+k} .$$

Therefore, since

$$S_n(v) = vX_1 + v \sum_{k=1}^{n-1} v^k X_{k+1} ,$$

$S_n(v)$ is seen to have the same distribution as $vX_1 + vS_{n-1}(v)$. Thus if $f_n(x)$ is the probability distribution function of $S_n(v)$, then the following convolution relationships will exist:

$$f_1(x) = f\left(\frac{x}{v}\right),$$

$$f_n(x) = \int_0^x f_{n-1}\left(\frac{x-y}{v}\right) f\left(\frac{y}{v}\right) dy \tag{10}$$

for $n = 2, 3, \ldots$ and $f(x)$ is the *pdf* of the X_k's.

Before deriving the integral equation for the *pdf* of $S_N(v)$, the following lemma is needed:

LEMMA 1. If X_k, $k = 1, 2, \ldots, n$ are *iid* random variables with finite mean, and the constants w_k are positive weights, let

$$Z_n = \sum_{k=1}^{n} w_k X_k \quad \text{and} \quad W_n = \sum_{k=1}^{n} w_k,$$

then for $k \in \{1, 2, \ldots, n\}$ and $n = 1, 2, \ldots$

$$E[X_k \mid Z_n = x] = \frac{x}{W_n}. \tag{11}$$

PROOF: By the symmetry of *iid* random variables and the fact that the weights are positive constants,

$$E[w_k X_k \mid Z_n = x] \propto w_k x.$$

Let π be the constant of proportionality. Summing both sides of the above expression yields

$$x = \pi W_n x,$$

i.e.,

$$\pi = \frac{1}{W_n}.$$

So

$$E[w_k X_k \mid Z_n = x] = \frac{w_k x}{W_n}$$

and equation (11) follows.

Q.E.D.

Consider the case where $w_k = v^k$ and $W_n = a_n$, then

$$\begin{aligned} E[X_1 \mid S_{n+1}(v) = x] &= \frac{x}{a_{n+1}} \\ &= \frac{1}{f_{n+1}(x)} \int_0^x \frac{y}{v} f_n\left(\frac{x-y}{v}\right) f\left(\frac{y}{v}\right) dy. \end{aligned} \tag{12}$$

We are now able to establish the main result of this paper.

THEOREM 1. Let $S_n(v)$ be defined as in equation (7) with *pdf* $g(x)$ for $x > 0$. If N has its probabilities satisfying the recursion in equation (8) and $\Sigma_{n=0}^{\infty} p_n = 1$, then for $x > 0$,

$$g(x) = p_1 f(x/v) + \int_0^x \left(a + \frac{by}{vx}\right) g\left(\frac{x-y}{v}\right) f(y/v)\, dy \tag{13}$$

with $\Pr[S_N(v) = 0] = p_0$.

PROOF: Since the X_k's are positive, $S_N(v) = 0$ if and only if $N = 0$. So $\Pr[S_N(v) = 0] = p_0$. For $x > 0$,

$$\begin{aligned}
g(x) &= \sum_{n=1}^{\infty} p_n f_n(x) \\
&= p_1 f_1(x) + \sum_{n=1}^{\infty} p_{n+1} f_{n+1}(x) \\
&= p_1 f(x/v) + \sum_{n=1}^{\infty} p_n \left(a + \frac{b}{a_{n+1}}\right) f_{n+1}(x) \\
&= p_1 f(x/v) + \sum_{n=1}^{\infty} a p_n \int_0^x f_n\left(\frac{x-y}{v}\right) f(y/v)\, dy + \\
&\quad + \sum_{n=1}^{\infty} p_n \frac{b}{a_{n+1}} f_{n+1}(x) \\
&= p_1 f(x/v) + \int_0^x a g\left(\frac{x-y}{v}\right) f(y/v)\, dy + \\
&\quad + \sum_{n=1}^{\infty} p_n \int_0^x \frac{by}{vx} f_n\left(\frac{x-y}{v}\right) f(y/v)\, dy \\
&= p_1 f(x/v) + \int_0^x \left(a + \frac{by}{vx}\right) g\left(\frac{x-y}{v}\right) f(y/v)\, dy
\end{aligned}$$

Q.E.D.

A similar result can be established if we assume that claims are subject to inflation at rate r and there is no interest. This can be accomplished by defining $w_k = (1+r)^k$, and using a new family of discrete claim number distributions with

$$p_n = p_{n-1}\left(a + \frac{b}{\ddot{s}_n}\right), \qquad \text{for} \qquad n = 1, 2, \ldots, \tag{14}$$

where

$$\ddot{s}_n = \sum_{k=1}^{n} (1+r)^k . \tag{15}$$

In this case

$$E[X_k \mid S_n(1+r) = x] = \frac{x}{\ddot{s}_n} . \tag{16}$$

The resulting integral equation is

$$g(x) = p_1 f(x/(1+r)) + \int_0^x \left(a + \frac{by}{(1+r)x}\right) g\left(\frac{x-y}{(1+r)}\right) f(y/(1+r))\, dy. \tag{17}$$

Note that in equation (13), for $0 < v < 1$, the argument of $g(.)$ in the integrand will exceed x, so $g(x)$ will depend on values of its argument between x and x/v. This will pose problems for obtaining numerical solutions. This problem does not arise in equation (17).

3. ANNUITY DISTRIBUTIONS

Equations (8) and (14) represent two new types of claim number distributions. However, they can be viewed as belonging to the same family of discrete annuity distributions because both equations can be written in the form:

$$p_n = p_{n-1}\left(a + \frac{b}{a(n, \delta)}\right), \qquad \text{for} \qquad n = 1, 2, \ldots, \tag{18}$$

where

$$a(n, \delta) = \sum_{k=1}^{n} e^{k\delta}, \qquad -\infty < \delta < \infty .$$

Here $\delta < 0$ can be viewed as the force of interest while $\delta > 0$ can be viewed as the force of inflation. This implies that from equation (9) and (15)

$$a(n, \delta) = \begin{cases} a_n & \text{if} \quad \delta < 0, \\ n & \text{if} \quad \delta = 0, \\ \ddot{s}_n & \text{if} \quad \delta > 0. \end{cases} \tag{19}$$

Thus the family of discrete distributions as described in equation (3) is a special case of the annuity distribution with $\delta = 0$.

For a non-defective annuity distribution to exist, its probabilties must sum to one, implying that

$$R(a, b, \delta) = 1 + \sum_{n=1}^{\infty} \prod_{k=1}^{n} \left(a + \frac{b}{a(k, \delta)}\right) \tag{20}$$

must coverge. There are several tests that can be used to check the convergence of $R(a, b, \delta)$, see MALIK (1984) or WILLMOT (1988). For example, the ratio-test ensures convergence if

$$\lim_{n\to\infty}\left(a+\frac{b}{a(n,\delta)}\right)=L<1\,.$$

Once $R(a, b, \delta)$ exists, the p_n's will be given by

$$p_n=\begin{cases}\dfrac{1}{R(a,b,\delta)} & \text{if} \quad n=0;\\[2ex] p_0\displaystyle\prod_{k=1}^{n}\left(a+\frac{b}{a(k,\delta)}\right) & \text{if} \quad n=1,2,3,\ldots\end{cases} \tag{21}$$

For given a and b that ensures the convergence of $R(a, b, \delta)$, one can easily evaluate the p_n's and the moments of the distribution. Unfortunately, closed form expressions are not easily obtainable these distributions, except of course when $\delta = 0$.

Further research is needed in the distributional properties of annuity distributions, the tail thickness, and the estimation of the parameters a and b. It will also be instructive to compare the various members of the family when $\delta = 0$ to those with the same parameters a and b but with $\delta \neq 0$. One would expect that the tails of these comparable distributions to become thicker as δ decreases.

REFERENCES

BOOGAERTS, P. and CRIJNS, V. (1987) Upper bounds on ruin probabilities in case of negative loadings and positive interest rates. *Insurance: Mathematics and Economics* **6**, 221–232.

BOOGAERT, P., HAEZENDONCK, J. and DELBAEN, F. (1988). Limit theorems for the present value of the surplus of an insurance portfolio. *Insurance: Mathematics and Economics* **7**, 131–138.

GARRIDO, J. (1988) Diffusion premiums for claim severities subject to inflation. *Insurance: Mathematics and Economics* **7**, 123–129.

GERBER, H. (1971) The discounted central limit theorem and its Berry-Esseen analogue. *Annals of Mathematical Statistics*, Vol. 42, **1**, 389–392.

MALIK, S. (1984) *Introduction to convergence*. Halstead Press, New York.

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

STRÖTER, B. (1985) The numerical evaluation of the aggregate claim density function via integral equations. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik* **17**, 1–14.

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

WATERS, H. (1983) Probability of ruin for a risk process with claim cost inflation. *Scandinavian Actuarial Journal* **66**, 148–164.

WILLMOT, G. (1988) Sundt and Jewell's family of discrete distributions. *ASTIN Bulletin* **18**, 17–29.