MATHEMATICAL FUN WITH THE COMPOUND BINOMIAL PROCESS

BY HANS U. GERBER

Université de Lausanne

ABSTRACT

The compound binomial model is a discrete time analogue (or approximation) of the compound Poisson model of classical risk theory. In this paper, several results are derived for the probability of ruin as well as for the joint distribution of the surpluses immediately before and at ruin. The starting point of the probabilistic arguments are two series of random variables with a surprisingly simple expectation (Theorem 1) and a more classical result of the theory of random walks (Theorem 2) that is best proved by a martingale argument.

KEYWORDS

Ruin; binomial model; compound binomial process; bonus malus; severity of ruin.

1. INTRODUCTION

The binomial model is a discrete time model, where the periodic premium is one and where in each period there is at most one claim. Such a model is of an independent interest, but it can also be used as an approximation to the continuous time compound Poisson model. In any case it has the advantage that the probabilistic reasoning is relatively elementary.

The purpose of this note is to adapt the methods and results of the continuous time compound Poisson model (see Gerber, 1988) to the compound binomial model.

2. THE MODEL

In this discrete time model

$$U_t = u + t - (X_1 + \cdots + X_{N_t}) \tag{1}$$

is the surplus of an insurance company at time t, $t = 0, 1, 2, \ldots$. Here $U_0 = u$ is the initial surplus, the premiums for each period are one, and N_t is the number of claims in the first t periods. It is assumed that this is a binomial process, i.e. that in any period there is a claim (with probability p) or no claim (with probability $q = 1 - p$), and that the occurrences of a claim in different periods are independent events. The amounts of these claims are denoted by $X_1, X_2, \ldots$. It is assumed that these random variables are independent and identically distributed and independent of the claim number process. We assume further that they are positive and integer valued; let

$$p(x) = Pr(X_i = x), \qquad x = 1, 2, 3, \ldots \tag{2}$$

denote their common probability function. Finally, we assume that the premiums contain a loading, i.e. that

$$p\mu < 1, \tag{3}$$

where μ is the mean claim size. In the following $u \geqslant 0$ is also integer valued.

'Ruin' is the event that $U_t \leqslant 0$ for some $t \geqslant 1$. We denote by $\psi(u)$ its probability, and let

$$T = \inf\{t \geqslant 1 : U_t \leqslant 0\} \tag{4}$$

be the period in which ruin occurs. By distinguishing according to what happens in the first period, and using the law of total probability, we see that

$$\psi(0) = q\psi(1) + p \tag{5}$$

and

$$\psi(u) = q\psi(u+1) + p \sum_{x=1}^{u} \psi(u+1-x)p(x) + p \sum_{x=u+1}^{\infty} p(x) \tag{6}$$

for $u = 1, 2, 3, \ldots$. We shall see that

$$\psi(0) = p\mu. \tag{7}$$

Thus formulae (5) and (6) can be used to calculate $\psi(u)$ recursively.

Illustrations: 1) If all claims are of size one, $p(1) = 1$ and $p(x) = 0$ for $x = 2, 3, \ldots$, ruin is only possible if $u = 0$ and if there is a claim in the first period. Thus

$$\psi(0) = p, \tag{8}$$

which confirms (7) and

$$\psi(u) = 0 \qquad \text{for } u = 1, 2, \ldots . \tag{9}$$

2) In the more interesting case where all claims are of size 2,

$$\begin{aligned} U_t &= u + t - 2N_t \\ &= u + (t - N_t) - N_t. \end{aligned} \tag{10}$$

This process and the resulting probability of ruin are discussed by many textbooks in the context of 'gambler's ruin problem'. Then

$$\psi(0) = 2p \tag{11}$$

and

$$\psi(u) = \left(\frac{p}{q}\right)^u \qquad \text{for } u = 1, 2, \ldots \tag{12}$$

is the solution of (7), (5) and (6).

In this note we shall try to show that there is more to the probability of ruin than its recursive calculation.

3. TWO FANCY SERIES AND THEIR SIMPLE VALUES

In the following let $S_0 = 0$ and $S_k = X_1 + \cdots + X_k$; we shall also use the notation $a^{(k)} = k!\binom{a}{k}$ for the factorial powers of a. The following strange result, which is the analogue of Theorem 1 of GERBER (1988), is the key to all further developments.

THEOREM 1.

a. For all x

$$\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{p}{q}\right)^k E[(S_k + x)^{(k)} q^{S_k + x}] = \frac{1}{1 - p\mu}.$$

b. For all $x \neq 0$

$$\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{p}{q}\right)^k E[(S_k + x - 1)^{(k-1)} q^{S_k + x}] = \frac{1}{x}.$$

PROOF: α. Let

$$g(z) = E[z^{X_i}] \tag{13}$$

denote the probability generating function of the X_i's. Then

$$E[(S_k + x)^{(k)} q^{S_k + x}] = q^k D^k [g(z)^k z^x] |_{z=q}, \tag{14}$$

where D is the derivative operator. Thus the left side in Theorem 1a is

$$\sum_{k=0}^{\infty} \frac{1}{k!} p^k D^k [g(z)^k z^x] |_{z=q}. \tag{15}$$

Using the Taylor series we obtain

$$\sum_{k=0}^{\infty} \frac{1}{k!} p^k D^k \sum_{j=0}^{\infty} \frac{1}{j!} (-p)^j D^j [g(z)^k z^x] |_{z=1}. \tag{16}$$

With the new summation variable $n = k + j$ this series can be written as

$$\sum_{n=0}^{\infty} \frac{1}{n!} p^n D^n \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} g(z)^k z^x |_{z=1}. \tag{17}$$

Because of the binomial formula this is

$$\sum_{n=0}^{\infty} \frac{1}{n!} p^n D^n [(g(z) - 1)^n z^x] |_{z=1}. \tag{18}$$

Since $g(1) = 1$ and $g'(1) = \mu$, this can be simplified to

$$\sum_{n=0}^{\infty} p^n \mu^n = \frac{1}{1 - p\mu}. \tag{19}$$

b. The left side in Theorem 1*b* is

$$\frac{1}{x}\, q^x + R, \tag{20}$$

where

$$\begin{aligned} R &= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\frac{p}{q}\right)^k E[(S_k + x - 1)^{(k-1)} q^{S_k + x}] \\ &= \sum_{k=1}^{\infty} \frac{1}{k!}\, p^k D^{k-1} [g(z)^k z^{x-1}]|_{z=q} \\ &= \sum_{k=1}^{\infty} \frac{1}{k!}\, p^k D^{k-1} \sum_{j=0}^{\infty} \frac{1}{j!} (-p)^j D^j [g(z)^k z^{x-1}]|_{z=1}. \end{aligned} \tag{21}$$

Setting $n = k + j$ we see that

$$\begin{aligned} R &= \sum_{n=1}^{\infty} \frac{1}{n!}\, p^n D^{n-1} \sum_{k=0}^{\infty} \binom{n}{k} (-1)^{n-k} g(z)^k z^{x-1}|_{z=1} \\ &\quad - \sum_{n=1}^{\infty} \frac{1}{n!} (-p)^n D^{n-1} z^{x-1}|_{z=1} \\ &= \sum_{n=1}^{\infty} \frac{1}{n!}\, p^n D^{n-1} [(g(z) - 1)^n z^{x-1}]|_{z=1} \\ &\quad - \frac{1}{x} \sum_{n=1}^{\infty} \frac{1}{n!} (-p)^n D^n z^x|_{z=1} \\ &= 0 - \frac{1}{x} (q^x - 1). \end{aligned} \tag{22}$$

Thus the left side of Theorem 1*b* is indeed $1/x$.

4. RESULTS FOR $u = 0$

In the following we shall assume $u = 0$. Some of the results can be easily translated to the more realistic situation where u is positive. For example $\psi(0)$, the probability of ruin with no initial surplus, can be interpreted more generally as the probability that the surplus will ever fall back on or below its initial level.

4.1. *The probability of ruin*

We shall derive formula (7). Since the process U_t is skipfree upwards and tends to infinity for $t \to \infty$, 'ruin' is equivalent to the event that there will be a visit at 0 ($U_t = 0$ for some $t \geqslant 1$). Thus the probability of ruin is the probability of a visit at 0.

Given S_k, a potential visit at 0 between (or with) the kth and the $(k+1)$th claim

https://doi.org/10.2143/AST.18.2.2014949 Published online by Cambridge University Press

takes place at time $t = S_k$, and such a visit will take place provided that $N_t = k$. Thus the conditional probability for a visit at 0 between (or with) the kth and the $(k+1)$th claim is

$$\Pr(N_t = k) = \binom{S_k}{k} p^k q^{S_k - k}. \tag{23}$$

Then the probability that the process visits 0 for the last time between (or with) the kth and $(k+1)$th claim is

$$E\left[\binom{S_k}{k} p^k q^{S_k - k}\right]\{1 - \psi(0)\}. \tag{24}$$

Hence the probability of ruin, which is also the probability of a visit at 0, is

$$\begin{aligned}\psi(0) &= \sum_{k=1}^{\infty} E\left[\binom{S_k}{k} p^k q^{S_k - k}\right]\{1 - \psi(0)\} \\ &= \sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{p}{q}\right)^k E[S_k^{(k)} q^{S_k}]\{1 - \psi(0)\}.\end{aligned} \tag{25}$$

From this and Theorem 1a it follows that

$$\psi(0) = \frac{p\mu}{1 - p\mu}\{1 - \psi(0)\}, \tag{26}$$

i.e. that $\psi(0) = p\mu$.

4.2. *The number of visits at a given level*

For $x \geqslant 0$ Theorem 1a has a natural probabilistic explication. Let us consider the number of visits at the level x. The probability that there are exactly n such visits is $\psi(0)^{n-1}\{1 - \psi(0)\}$, for $n = 1, 2, \ldots$. Thus the expected number of visits at x is $1/(1 - \psi(0))$, which is the expression on the right side of Theorem 1a.

The expected number of visits at x can be calculated in a different way. Given S_k, there will be a visit at x between (or with) the kth claim and the $(k+1)$th claim at time $t = S_k + x$, provided that $N_t = k$. Thus the probability for a visit at x between (or with) the kth claim and the $(k+1)$th claim is

$$E\left[\binom{S_k + x}{k} p^k q^{S_k + x - k}\right]. \tag{27}$$

Summation over k $(k = 0, 1, 2, \ldots)$ gives the expected number of visits at x; but this is the expression on the left side of Theorem 1a.

4.3. *A classical result*

The following result, due to Dwass, Dinges and Keilson is about two conditional probabilities given the event $A = \{U_t = x \text{ and } N_t = k\}$, where x is a positive integer.

THEOREM 2:

a) $\Pr(U_\tau < x \text{ for } \tau = 1, \ldots, t-1 \mid A) = x/t$,
b) $\Pr(U_\tau > 0 \text{ for } \tau = 1, \ldots, t-1 \mid A) = x/t$.

An elegant proof due essentially to DELBAEN and HAEZENDONCK (1985) is as follows.

a) We consider the sequence

$$M_\tau = \frac{x - U_\tau}{t - \tau} \quad \text{for} \quad \tau = 0, 1, \ldots, t-1. \tag{28}$$

Given A, this is a martingale. Thus if ν is a stopping time, it follows that

$$M_0 = E[M_\nu \mid A]. \tag{29}$$

Now let ν be the first time when $M_\tau = 0$ ($U_\tau = x$), or $\nu = t-1$ if $M_\tau > 0$ ($U_\tau < x$) for $\tau = 1, \ldots, t-1$. For this stopping time (29) yields Theorem 2a).

b) We consider the sequence

$$M_\tau = \frac{U_\tau}{\tau} \quad \text{for} \quad \tau = 1, \ldots, t. \tag{30}$$

Given A, this is a backward martingale. Thus, if ν is a backward stopping time, it follows that

$$M_t = E[M_\nu \mid A]. \tag{31}$$

Now let ν be the last time τ when $M_\tau = U_\tau = 0$, or $\nu = 1$ if $M_\tau > 0$ for $\tau = 1, 2, \ldots, t-1$. For this backward stopping time (31) yields Theorem 2b).

4.4. *The probability of reaching a given level*

Let x be a positive integer. By Theorem 2a the conditional probability (given S_k) that the level x is visited for the first time between (or with) the kth claim and the $(k+1)$th claim is

$$\binom{t}{k} p^k q^{t-k} \frac{x}{t} = x \frac{1}{k!} \left(\frac{p}{q}\right)^k (t-1)^{(k-1)} q^t, \tag{32}$$

where $t = S_k + x$. Thus the probability that the level x will ever be visited is

$$x \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{p}{q}\right)^k E[(S_k + x - 1)^{(k-1)} q^{S_k + x}]. \tag{33}$$

Of course this probability is one, which illustrates Theorem 1*b*.

4.5. *The surplus immediately before and at ruin*

For $x = 1, 2, 3, \ldots$ and $y = 0, 1, 2, \ldots$ we can use Theorem 2b to see that the conditional probability (given S_k) for ruin with the $(k+1)$th claim, such that

$U_{T-1} = x$ and $U_T = -y$, is

$$(34) \quad \binom{S_k + x}{k} p^k q^{S_k + x - k} \frac{x}{S_k + x} pp(x+1+y)$$

$$= \frac{1}{k!} \left(\frac{p}{q}\right)^k (S_k + x - 1)^{(k-1)} q^{S_k + x} xpp(x+1+y).$$

We take expectations, sum over k and use Theorem 1b to obtain

$$(35) \qquad \Pr(T < \infty, U_{T-1} = x, U_T = -y) = pp(x+1+y),$$

valid for $x = 1, 2, 3, \ldots$ and $y = 0, 1, 2, \ldots$. This formula is also true for $x = 0$ (then it gives simply the probability that there is a claim of size $1 + y$ in the first period). The compound Poisson analogue of (35) is derived by DUFRESNE and GERBER (1988).

From (35) we get

$$(36) \qquad \Pr(T < \infty, U_T = -y) = p \sum_{x=0}^{\infty} p(x+1+y) = p[1 - P(y)].$$

This is the discrete analogue of Theorem 12.4 of BOWERS *et al.* (1987). Finally, we may sum (36) over y to confirm (7) once again.

5. RESULTS FOR $u > 0$

The reasoning of Section 4.1 can be adapted to this more general situation. Given S_k, a visit at 0 between (or with) the kth claim and the $(k+1)$th claim is possible if $S_k \geqslant u$; such a visit will take place at time $t = S_k - u$, provided that $N_t = k$. Thus the conditional probability for a last visit at 0 between (or with) the kth claim and the $(k+1)$th claim is

$$(37) \qquad \binom{S_k - u}{k} p^k q^{S_k - u - k} \{1 - \psi(0)\}$$

if $S_k \geqslant u$, and 0 otherwise; the expression

$$(38) \qquad \binom{(S_k - u)_+}{k} p^k q^{S_k - u - k} \{1 - \psi(0)\}$$

is the conditional probability for both cases. Then the probability for a last visit at 0 before (or with) the kth claim and the $(k+1)$th claim is

$$(39) \qquad E\left[\binom{(S_k - u)_+}{k} p^k q^{S_k - u - k}\right] \{1 - \psi(0)\}.$$

Hence

$$(40) \qquad \psi(u) = \sum_{k=1}^{\infty} E\left[\binom{(S_k - u)_+}{k} p^k q^{S_k - u - k}\right] \{1 - \psi(0)]$$

$$= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\frac{p}{q}\right)^k E[(S_k - u)_+^{(k)} q^{S_k - u}] \{1 - p\mu\}.$$

There is an alternative series expression for $\psi(u)$. From Theorem 1a it follows that

$$(41) \qquad \sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{p}{q}\right)^k E[(S_k-u)^{(k)} q^{S_k-u}] = \frac{1}{1-p\mu} - q^{-u}.$$

Combining (40) with (41) we obtain

$$(42) \qquad \frac{1-\psi(u)}{1-p\mu} = q^{-u} + \sum_{k=1}^{u-1} \frac{1}{k!}\left(\frac{p}{q}\right)^k E[(S_k-u)^{(k)} q^{S_k-u} I_{[S_k<u]}]$$

This is the discrete time analogue of Shiu's formula (1988). Note that in (42), unlike in (40), the summation with respect to k and 'E' is finite for any value of u.

ILLUSTRATION. If all claims are of constant size m (a positive integer), it follows that $S_k = km$, and from (42) that

$$(43) \qquad \frac{1-\psi(u)}{1-pm} = q^{-u} + \sum_{k=1}^{[(u-1)/m]} \binom{km-u}{k}\left(\frac{p}{q}\right)^k q^{km-u}.$$

This result generalizes formulas (8), (9), (11) and (12), which are for $m=1$ and $m=2$. It is the discrete analogue of a well known formula in classical ruin theory, see, e.g. FELLER (1966, formula (2.11) of chapter XIV.2). DUFRESNE (1988) shows that there is a close connection between the probability of ruin and the stationary distribution of a bonus-malus system; thus (43) can be used to describe such a stationary distribution analytically.

REFERENCES

BOWERS, N. L., GERBER, H. U., HICKMAN, J. C., JONES D. A., and NESBITT, C. J. (1987). *Actuarial Mathematics*. Society of Actuaries, Itasca, Illinois, U.S.A.

DELBAEN, F. and HAEZENDONCK, J. (1985). Inversed martingales in risk theory. *Insurance: Mathematics and Economics* **4**, 201–206.

DUFRESNE, F. (1988). Distributions stationnaires d'un système bonus-malus et probabilité de ruine. *ASTIN Bulletin* **18**, 31–46.

DUFRESNE, F. and GERBER, H. U. (1988). The surpluses immediately before and at ruin, and the amount of the claim causing ruin. Forthcoming in *Insurance: Mathematics and Economics*.

FELLER, W. (1966). *An Introduction to Probability Theory and its Applications, Volume 2.*, Wiley, New York.

GERBER, H. U. (1988). Mathematical Fun with Risk Theory. *Insurance: Mathematics and Economics* **7**, 15–23.

SHIU, E. S. W. (1988). Calculation of the probability of eventual ruin by Beekman's convolution series. *Insurance: Mathematics and Economics* **7**, 41–47.

HANS U. GERBER

Ecole des H.E.C., Université de Lausanne, 1015 Lausanne, Switzerland