A NOTE ON RANDOM SURVIVORSHIP GROUP BENEFITS

By Colin M. Ramsay

University of Nebraska, 310 Burnett Hall
Lincoln, NE USA 68588-0307 (402)-472-5823

Abstract

Consider a group of n independent lives age x where each life puts § 1 in a fund at time 0. The fund earns interest at rate i, and at the end of t years the accumulated value of the fund is divided equally among the survivors. The traditional approach to calculating the expected lump sum benefit per survivor from the initial group of n lives is based on the concept of a deterministic survivorship group. This approach ignores the stochastic nature of the survivorship process. In reality, the benefit per survivor is actually a random variable with an expected value which depends on the first inverse moment of a positive binomial random variable. Using GRAB's and SAVAGE's (1954) recursive formula for the first inverse moment, it is shown that the traditional approach yields a fairly accurate approximation to the solution even when one assumes a random number of survivors.

Keywords

Inverse moments; positive binomial distribution; actuarial accumulated value.

1. THE PROBLEM

Consider two well known problems in actuarial life contingencies:

Problem (1): Each member of a group of n independent lives age x makes a single payment of 1 into a fund at time 0. At time t the accumulated value of the fund is divided equally among the survivors, if any. Assuming the fund earns interest at rate i per annum and $_tp_x$ is the probability of (x) surviving t years, calculate the expected share of each survivor. The traditional answer of this problem is

$$(1) \qquad \text{Expected Share} = \frac{(1+i)^t}{_tp_x}$$

Problem (2): Each member of a group of n independent lives age x deposits an amount 1 into a fund at the start of each year as long as he/she is alive. At time t,

the accumulated value of the fund is divided equally among the survivors, if any. If the fund earned interest at rate i per annum, calculate the expected share of each survivor. The traditional solution is

(2)   Expected Share $= \ddot{s}_{x:\overline{t}|}$

These problems are often discussed when the concept of the actuarial accumulated value is introduced. See, for example BOWERS et al. (1986, Chapter 5) or NEILL (1977, Chapter 2.3). Fundamental to the derivation of the traditional "solutions" given in equations (1) and (2) is the notion of a deterministic survivorship group; see BOWERS et al. (1986, Chapter 3.4). In a deterministic survivorship group, deaths are not random and the proportion of survivors at any age is always known exactly. Constrasting this is a random survivorship group (BOWERS et al., Chapter 3.3) where the number of survivors at any age is random variable.

Since, in any practical situation, the number of survivors will be a random variable, it is clear that the traditional solutions can only be viewed as approximate solutions. To derive the true solutions to problems (1) and (2), the actual amount of money each survivor receives must be expressed as random variables, then their expectations must be found. This can be accomplished as follows: Let S_1 and S_2 be the actual share of the fund that each survivor receives in problems (1) and (2) respectively. If there are no survivors then $S_1 = S_2 = 0$ so for problem (1),

(3) $S_1 = \begin{cases} n(1+i)^t/N & \text{if} \quad N = 1, 2, \ldots, n \\ 0 & \text{if} \quad N = 0 \end{cases}$

and for problem (2)

(4) $S_2 = \begin{cases} (N\ddot{a}_{\overline{t}|} + \sum_{i=1}^{D} \ddot{a}_{\overline{K_i+1|}})(1+i)^t/N & \text{if} \quad N = 1, 2, \ldots, n \\ 0 & \text{if} \quad N = 0 \end{cases}$

where

N = number of survivors at age $x+t$ from the group of n lives age x.

D = $n - N$ is the number of deaths (from the initial n lives) between ages x and $x+t$.

K_i = curtate future life time for the ith life who died.

By assumption, the n lives are mutually independent so N has a binomial distribution with parameters n and $_t p_x$. From equations (3) and (4), both $E[S_1]$ and $E[S_2]$ depend on $E[N^{-1}|N > 0]$, i.e., the first inverse moment of the positive binomial distribution. More on this in Section (2).

It should be pointed out that, in addition to depending on N, S_2 also depends on the random variable K_i, for $i = 1, 2, \ldots, D$, so expectations must

be taken over the K_i's as well. To this end, one can write

$$E[S_2] = E[E[S_2|N]]$$

$$= \sum_{m=1}^{n} E[S_2|N = m] \Pr [N = m]$$

$$= \sum_{m=1}^{n} \left(\frac{m \, \ddot{a}_{\overline{t}|} + \sum_{i=1}^{n-m} E[\ddot{a}_{\overline{K_i+1}|}]}{m} \right) (1+i)^t \Pr [N = m].$$

But given $N = m$, there are $n - m$ deaths, each with $K_i < t$. For convenience it is assumed that t is an integer, i.e., $t = 1, 2, \ldots$.

(5)
$$E[\ddot{a}_{\overline{K_i+1}|}|K_i < t] = \sum_{k=0}^{t-1} {}_{k|}q_x \, \ddot{a}_{\overline{k+1}|}/{}_tq_x$$

$$= (\ddot{a}_{x:\overline{t}|} - {}_tp_x \, \ddot{a}_{\overline{t}|})/{}_tq_x$$

where ${}_tq_x = 1 - {}_tp_x$. It follows that

(6)
$$E[S_2] = (1+i)^t \sum_{m=1}^{n} \left[\ddot{a}_{\overline{t}|} + \frac{(n-m)}{m} \frac{(\ddot{a}_{x:\overline{t}|} - {}_tp_x \, \ddot{a}_{\overline{t}|})}{(1 - {}_tp_x)} \right] \Pr [N = m]$$

$$= \sum_{m=1}^{n} \left[\ddot{s}_{\overline{t}|} + \left(\frac{n}{m} - 1 \right) (\ddot{s}_{x:\overline{t}|} - \ddot{s}_{\overline{t}|}) \frac{{}_tp_x}{{}_tq_x} \right] \Pr [N = m].$$

Let ${}_tc_x^{(n)}$ denote the sum

(7)
$${}_tc_x^{(n)} = \sum_{m=1}^{n} \frac{n}{m} \Pr [N = m]$$

$$= \sum_{m=1}^{n} \frac{n}{m} \binom{n}{m} ({}_tp_x)^m ({}_tq_x)^{n-m}.$$

Thus equations (3) and (6) can be written as

(8)
$$E[S_1] = (1+i)^t \, {}_tc_x^{(n)}$$

(9)
$$E[S_2] = \left[\ddot{s}_{\overline{t}|}(1 - ({}_tq_x)^n) + ({}_tc_x^{(n)} - (1 - ({}_tq_x)^n)) (\ddot{s}_{x:\overline{t}|} - \ddot{s}_{\overline{t}|}) \frac{{}_tp_x}{{}_tq_x} \right].$$

Using equation (7), one can compute ${}_tc_x^{(n)}$ directly for any finite n. Unfortunately, for large n, the bionomial probabilities are cumbersome to find. However, LING (1992) provided an exact method of calculating the cumulative distribution function of a binomial distribution via the F-distribution. This method requires that the F-probabilities must be available. Ling also provides references for accurate approximations to binomial probabilities.

In view of the difficulties associated with directly evaluating the summation on the right hand side of equation (2), different approaches are needed to evaluate $_t c_x^{(n)}$. In the next section, it will be proved that $_t c_x^{(n)}$ is actually proportional to the first inverse moment of the corresponding positive binomial distribution. The recurrence relationship between the successive first inverse moments provided by GRAB and SAVAGE (1954) is used to evaluate $_t c_x^{(n)}$.

2. THE POSITIVE BINOMIAL DISTRIBUTION

Suppose a random variable X has binomial distribution with parameters n and p, i.e.,

$$f_m(n, p) = \Pr [X = m] = \binom{n}{m} p^m q^{n-m}, \qquad m = 0, 1, \ldots, n$$

where $p + q = 1$ and $0 \le p \le 1$. Then a random variable Y is called a *positive binomial random variable* if its probabilities satisfy

$$(10) \qquad \Pr [Y = m] = \frac{f_m(n, p)}{1 - q^n}, \qquad m = 1, 2, \ldots, n.$$

The inverse moments of positive binomial random variables have been studied by several authors including GRAB and SAVAGE (1954) who provided a recurrence relation for the first inverse moment. Let $b_k(n, p) = E[Y^{-k}]$ be the kth inverse moment of Y. Grab and Savage proved that

$$(11) \qquad b_1(n+1, p) = \frac{q(1-q^n)}{(1-q^{n+1})} b_1(n, p) + \frac{1}{n+1}, \qquad n = 0, 1, \ldots$$

with $b_1(n, p) = 0$. MENDENHALL and LEHMAN (1960) used the beta distribution to approximate $b_k(n, p)$. In particular, Mendenhall and Lehman suggested the approximation

$$(12) \qquad b_k(n, p) = \frac{1}{n^k} \prod_{r=1}^{k} \frac{(a+b-r)}{(a-r)}$$

where $a = (n-1)p$ and $b = (n-1)q$.

If one is interested in computing higher moments, an exact recurrence relatoinship among the higher order inverse moments of Y is given in GOVINDARAJULU (1963). For example, Govindarajulu proved that the second inverse moments satisfy

$$(13) \qquad b_2(n+1, p) = \frac{q(1-q^n)}{(1-q^{n+1})} \left(\frac{b_1(n, p)}{n+1} + b_2(n, p) \right) + \frac{1}{(n+1)^2}$$

with $b_2(0, p) = 0$. So $b_1(b, p)$ and $b_2(n, p)$ can jointly be evaluated using the recursive expressions in equations (11) and (13). Note that equation (13) will not be used in the sequel.

WALKER (1984) actually suggested three approximations to $b_k(n, p)$. Let $d = 1 + b$ and $e = ac$, then the approximations are as follows:

(14) $$b_k(n, p) = c^{-k} \prod_{r=1}^{k} \left(\frac{a+b-r}{a-r} \right)$$

where, for approximation:

(i) $c = n + (n-1)q(2p-1)$, $d = 2(n-1)p$ and $e = 2n(n-1)p^2$;

(ii) $c = n$, $d = (n-1)(1-q^n)/V$ and $e = n(n-1)p/V$ where
$V = 1 - q^n - npq^{n-1}$; and finally

(iii) $c = n + (n-1)q(2p-1) + 2n(n-1)^2 p^3/W$

$d = 2(n-1)p(1-q^n)/W$ and $e = 2n(n-1)p^2/W$ where

$W = 1 - q^n - npq^{n-1}(q+np)$

When $np > 10$ the approximations cited above are all reasonably accurate. However, since equation (11) is exact, it will be used to calculate $E[S_1]$ and $E[S_2]$.

3. TRADITIONAL VS EXACT EXPECTED SHARES

In this section, the traditional answers given in equations (1) and (2) are compared to the exact answers given in equations (8) and (9). Equation (11) will be used to compute $_t c_x^{(n)}$. Note that

(15) $$_t c_x^{(n)} = n(1 - (_t q_x)^n) b_1(n, \, _t p_x)$$

Let ρ_i be the ratio of $E[S_i]$ to the traditional approximation to this expectation, i.e.,

(16) $$\rho_1 = \frac{E[S_1]}{(1+i)^t /_t p_x}$$

$$= \, _t c_x^{(n)} \, _t p_x$$

(17) $$\rho_2 = \frac{E[S_2]}{\ddot{s}_{x:\overline{t}|}} \ .$$

Mortality is assumed to follow that given in BOWERS et al. (page 72, equation (3.7.1)), i.e.,

$$\mu_x = A + Bc^x$$

where $A = 0.0007$, $B = 0.00005$ and $c = 10^{0.04}$. In addition, it is assumed that $x + t = 65$, i.e., a person must survive to age 65 in order to receive survival benefits. For various ages $x = 20, 30, \ldots, 60$, the ratios ρ_1 and ρ_2 are computed for different groups sizes. After using several different interest rates, the ratio ρ_2 appears to be fairly insensitive to interest rate changes, so only the ratios for

COLIN M. RAMSAY

$i = 6\%$ are given. Note that ρ_1 is independent of the rate of interest i. Group sizes from 1 to 500 are used with $b_1(n, p)$ calculated using equation (11).

From Tables 1 and 2, it is clear that for $n \geq 10$, the ratios are very close to 1 so the traditional approach is fairly accurate even though it uses a deterministic approach.

TABLE 1

ρ_1 WHEN (x) SURVIVES TO AGE 65

n	$x = 20$	$x = 30$	$x = 40$	$x = 50$	$x = 60$
5	1.08105	1.07628	1.06849	1.05344	1.02311
10	1.03317	1.03117	1.02798	1.02197	1.00986
20	1.01503	1.01416	1.01277	1.01011	1.00461
30	1.00973	1.00918	1.00828	1.00657	1.00301
40	1.00720	1.00679	1.00613	1.00487	1.00224
50	1.00571	1.00539	1.00487	1.00387	1.00178
100	1.00281	1.00265	1.00240	1.00191	1.00088
200	1.00139	1.00132	1.00119	1.00095	1.00044
300	1.00093	1.00087	1.00079	1.00063	1.00029
400	1.00069	1.00066	1.00059	1.00047	1.00022
500	1.00055	1.00052	1.00047	1.00038	1.00017

TABLE 2

ρ_2 WHEN (x) SURVIVES TO AGE 65

n	$x = 20$	$x = 30$	$x = 40$	$x = 50$	$x = 60$
5	1.07305	1.06573	1.05502	1.03878	1.01547
10	1.02991	1.02688	1.02250	1.01596	1.00661
20	1.01355	1.01221	1.01027	1.00734	1.00309
30	1.00878	1.00792	1.00666	1.00477	1.00202
40	1.00649	1.00586	1.00493	1.00354	1.00150
50	1.00515	1.00465	1.00391	1.00281	1.00119
100	1.00253	1.00229	1.00193	1.00139	1.00059
200	1.00126	1.00113	1.00096	1.00069	1.00029
300	1.00084	1.00075	1.00064	1.00046	1.00019
400	1.00063	1.00057	1.00048	1.00034	1.00015
500	1.00050	1.00045	1.00038	1.00027	1.00012

Tables 3 and 4 show that ratios based on the Mendenhall-Lehman approximation to $b_1(n,p)$, i.e.,

(18) $$b_1(n, p) = \frac{(n-2)}{n(a-1)} \quad \text{with} \quad a = (n-1)p.$$

Substituting this approximation into equation (15) results in

(19) $$_t c_x^{(n)} \approx (1 - {}_t q_x^n)(n-2)/(a-1).$$

If the expected number of survivors $(n \times {}_t p_x)$ is large, the approximation given in equation (19) becomes fairly accurate and may be used (instead of equation (11)) to estimate $_t c_x^{(n)}$.

TABLE 3

ρ_1 WHEN (x) SURVIVES TO AGE 65

n	$x = 20$	$x = 30$	$x = 40$	$x = 50$	$x = 60$
5	1.10104	1.09493	1.08517	1.06678	1.02980
10	1.03581	1.03374	1.03042	1.02408	1.01097
20	1.01561	1.01472	1.01329	1.01056	1.00485
30	1.00998	1.00941	1.00851	1.00676	1.00311
40	1.00733	1.00692	1.00625	1.00497	1.00229
50	1.00580	1.00547	1.00494	1.00393	1.00181
100	1.00283	1.00267	1.00242	1.00192	1.00089
200	1.00140	1.00132	1.00119	1.00095	1.00044
300	1.00093	1.00088	1.00079	1.00063	1.00029
400	1.00070	1.00066	1.00059	1.00047	1.00022
500	1.00056	1.00052	1.00047	1.00038	1.00017

TABLE 4

ρ_2 WHEN (x) SURVIVES TO AGE 65

n	$x = 20$	$x = 30$	$x = 40$	$x = 50$	$x = 60$
5	1.09107	1.08182	1.06843	1.04847	1.01995
10	1.03230	1.02910	1.02446	1.01748	1.00735
20	1.01407	1.01270	1.01069	1.00767	1.00325
30	1.00900	1.00812	1.00684	1.00491	1.00208
40	1.00661	1.00597	1.00503	1.00361	1.00153
50	1.00523	1.00472	1.00398	1.00286	1.00121
100	1.00255	1.00230	1.00194	1.00140	1.00059
200	1.00126	1.00114	1.00096	1.00069	1.00029
300	1.00084	1.00076	1.00064	1.00046	1.00020
400	1.00063	1.00057	1.00048	1.00034	1.00015
500	1.00050	1.00045	1.00038	1.00027	1.00012

I was very surprised to see that the traditional actuarial approach provided such a good approximation even when there is actually a random survivorship group situation. It appears that there may be no need to use the exact theory unless the group is very small, say less than 20 lives.

Note the apparent result, from the tables above, that the exact expected benefit per survivor is greater than that suggested by the traditional approach, i.e., $\rho_1 > 1$ and $\rho_2 > 1$. In the case of ρ_1, this is due to the Schwartz inequality (as Grab and Savage noted in their equation (7)), where

$$E\left[\frac{1}{N}\right] \geq \frac{1}{E[N]}.$$

The explanation for $\rho_2 > 1$ is not as obvious. However these results make sense from the point of view of risk and return. Since the deterministic survivorship group is riskless, the expected return to the "survivors" must be less than the expected return demanded by those in the random survivorship group.

REFERENCES

BOWERS, N.L., GERBER, H.U., HICKMAN, J.C., JONES, D.A. and NESBITT, C.J. (1986) *Actuarial Mathematics*. Society of Actuaries, Itasca, IL.

GOVINDARAJULU, Z. (1963) "Recurrence relations for the inverse moments of the positive binomial variable". *Journal of the American Statistical Association* **58**, 468–473.

GRAB, E.L. and SAVAGE, I.R. (1954) "Tables of the expected value of 1/X for positive Bernouilli and Poisson variables". *Journal of the American Statistical Association* **49**, 169–177.

LING, R.F. (1992) "Just say no to binomial (and other discrete distributions) Tables". *The American Statistician* **46**, No. 1, 53–54.

MENDENHALL, W. and LEHMAN, E.H. Jr. (1960) "An approximation to the negative moments of the positive binomial useful in life testing". *Technometrics* **2**, No. 2, 227–242.

NEILL, A. (1977) *Life Contingencies*. Heinemann, London.

WALKER, J.J. (1984) "Improved approximations for the inverse moments of the positive binomial distribution". *Communications in Statistics: Simulations and Computations* **13**, No. 4, 515–519.

COLIN M. RAMSAY
University of Nebraska, 310 Burnett Hall,
Lincoln, NE, USA 68588-0307, (402)-472-5823.