Improved approximations for the distribution of a heterogeneous risk portfolio

Autor(en): **Jewell, William S. / Sundt, Bjørn**

Objekttyp: **Article**

Zeitschrift: **Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries**

Band (Jahr): **- (1981)**

Heft 2

PDF erstellt am: **22.07.2024**

Persistenter Link: https://doi.org/10.5169/seals-551242

William S. Jewell, Berkeley, and Bjørn Sundt, Oslo/Zurich

Improved Approximations for the Distribution of a Heterogeneous Risk Portfolio

Introduction

Let $(\tilde{x}_i; i = 1, 2, \dots N)$ be a fixed set of independent, non-identically distributed, integer-valued random variables for which the probability that any $\tilde{x}_i = 0$ is significant; we wish to find the distribution of the sum $\tilde{y} = x_1 + \tilde{x}_2 + \dots + \tilde{x}_N$. In principle, the discrete density of $\tilde{y}$ is calculated as the N-fold convolution of the discrete densities of the individual $\tilde{x}_i$; however, this task is already very time-consuming on digital computers for N larger than, say, 1,000, if the $\tilde{x}_i$ take on more than a few different values.

An approximate method, used for many years by actuaries, utilizes the fact that many terms in the sum may have value zero and computes $\tilde{y}$ as if it were the sum of a *random* number of independent identically distributed random variables; in this method, the first moment of $\tilde{y}$ is matched exactly, and the second moment is matched approximately.

In this paper, we present improved approximations that provide a much closer fit to the second moment, yet maintain a simple, recursive algorithm for computing the density of the random sum. Limited computational experience indicates that these approximations to the distribution and other functions of $\tilde{y}$ are much closer to their true values than in the classical method.

The Heterogeneous Portfolio

For the moment, we assume that the $\tilde{x}_i$ take only non-negative values in the range $[0, 1, \dots R]$, and we separate the given discrete density of $\tilde{x}_i$ as follows:

$$p_i \quad = \Pr\{\tilde{x}_i = 0\} = 1 - q_i, \qquad (i = 1, 2, \dots N) \tag{1}$$

$$f_i(x) = \Pr\{\tilde{x}_i = x \mid x_i > 0\}. \qquad (x = 1, 2, \dots R). \tag{2}$$

(This is a traditional notation.)

We wish to calculate the discrete density g of the sum

$$\tilde{y} = \sum_{i=1}^{N} \tilde{x}_i, \qquad (y = 0, 1, \ldots NR) \tag{3}$$

which is given exactly by the N-fold discrete convolution:

$$g(y) = \mathop{\ast}_{i=1}^{N} [p_i\, \delta(y) + q_i f_i(y)], \tag{4}$$

where $\delta(y) = 1$ if $y = 0$, and is zero otherwise. The approximation to be described requires that "most" of the p_i be "rather large".

Denote the first two moments of the positive part of the random variables by

$$m_i = \boldsymbol{E}\{\tilde{x}_i | \tilde{x}_i > 0\} = \Sigma x f_i(x);\ v_i = \boldsymbol{V}\{\tilde{x}_i | \tilde{x}_i > 0\} = \Sigma (x - m_i)^2 f_i(x). \tag{5}$$

Then it is easy to show that the first two moments of the sum y are:

$$\boldsymbol{E}(\tilde{y}) = \sum_{i=1}^{N} q_i m_i; \tag{6}$$

$$\boldsymbol{V}(\tilde{y}) = \sum_{i=1}^{N} q_i v_i + \sum_{i=1}^{N} p_i q_i (m_i)^2. \tag{7}$$

The evaluation of $g(y)$ is often required for insurance *risk portfolios,* where $i = 1, 2, \ldots N$ indexes the policies in the portfolio, assumed independent; p_i is the usually significant *no-claim probability* during a certain *exposure period*; q_i is the probability of at least one claim; and $f_i(x)$ is the density of aggregate claims during the exposure period for policy i, given that at least one claim occurs.

The situation is particularly simple in life insurance, as usually just one claim occurs at death, and the $f_i(x)$ are often only one- or two-point densities (e.g., the face value of a policy i payable at the death of the assured, who has mortality rate q_i in this exposure period). Often, only the q_i change from one exposure period to the next. Approximation methods have become less important in such simple cases, especially with N small, as modern computers can often calculate the exact convolution (4) directly. However, for large portfolios with arbitrary $f_i(x)$, the problem of approximating $g(y)$ still remains. Most approaches have been based upon moment-matching, using (6) and (7).

The Collective Risk Model as an Approximation

One useful idea, from both the theoretical and computational points of view, is to approximate the inhomogeneous, fixed portfolio by a homogeneous *risk collective,* in which we replace the individual policies by a mass of similar, anonymous policies, and assume that $\tilde{y}$ is the sum of a *random number,* $\tilde{n}$, of independent claims that are *identically distributed samples,* $(w_1, w_2, \ldots w_{\tilde{n}})$ of a positive random variable, $\tilde{w}$, with some *prototypical claim density,* $f(w)$. If

$$\pi_n = \Pr\{\tilde{n} = n\}, \ (n = 0, 1, \ldots); f(w) = \Pr\{\tilde{w} = w\}, \ (w = 1, 2, \ldots); \qquad (8)$$

then this leads to the well-known *compound law* of risk theory:

$$g(y) = \pi_0 \delta(y) + \sum_{n=1}^{\infty} \pi_n [f(y)]^{n*}. \qquad (9)$$

The rationale for this approximation is easily seen. If the p_i are significant, then the sum $\tilde{y} = \tilde{x}_1 + \tilde{x}_2 + \ldots + \tilde{x}_N$ will have a varying number of non-zero terms; the sum could then be represented by $\tilde{y} = \tilde{w}_1 + \tilde{w}_2 + \ldots + \tilde{w}_{\tilde{n}}$, where these all-positive terms could be considered to be identically-distributed samples from some "representative" claim density, calculated by weighting each $f_i(x)$ according to its probability of occurence, q_i.

If the prototypical claim moments are:

$$m = \boldsymbol{E}(\tilde{w}) = \Sigma w f(w); \ v = \boldsymbol{V}(\tilde{w}) = \Sigma (w-m)^2 f(w), \qquad (10)$$

then the moments of the random sum (9) will be:

$$\boldsymbol{E}(\tilde{y}) = \boldsymbol{E}(\tilde{n}) m, \qquad (11)$$

$$\boldsymbol{V}(\tilde{y}) = \boldsymbol{E}(\tilde{n}) v + \boldsymbol{V}(\tilde{n}) m^2. \qquad (12)$$

For a good approximation, the moments (11) (12) must be matched as closely as possible to the exact values (6) (7), so that $g(y)$ and related functions calculated via (9) will match values calculated via (4).

We are, of course, free in devising an approximation to choose π_n and $f(w)$ in any way we choose. But the most natural way to fix the prototypical claim density, consistent with the risk theory interpretation, is as the weighted sum:

$$f(w) = \frac{\Sigma q_i f_i(w)}{\Sigma q_i}, \qquad (w = 1, 2, \ldots R). \qquad (13)$$

(This choice is also invariant under pre-aggregation of the policies in a consistent way, for example, by lumping together all policies with the same single face value and adding their q_i's). With this choice, the moments of $\tilde{w}$ become:

$$m = (\Sigma q_i m_i)/(\Sigma q_j); \tag{14}$$

$$v + m^2 = \left[\Sigma q_i v_i + \Sigma q_i (m_i)^2\right]/(\Sigma q_j). \tag{15}$$

If (11) and (12) are to be matched *exactly* to (6) (7), then this implies that the counting density, π_n, must be chosen so that:

$$\boldsymbol{E}(\tilde{n}) = \Sigma q_j, \tag{16}$$

and

$$\boldsymbol{V}(\tilde{n}) = \Sigma q_j - \Sigma q_i^2 (m_i/m)^2. \tag{17}$$

The mean of $\tilde{n}$ is just the mean number of positive terms in (3); however, the variance of counts is *not* the variance of the number of positive terms, $\Sigma q_j p_j$, unless the policies have identical face values. This is because we are matching moments between two different models, one where the sampling is without replacement, and another where the sampling is independent.
Note that, in certain unusual cases where the p_i are small and the m_i are quite different from another, $V(\tilde{n})$ in (17) may be negative; in other words, the approximation cannot be used. For instance, if $N = 2$, $m_1 = 1$, $m_2 = 7$, and $q_1 = q_2 = q$, then we find that q must be smaller than 0.64 to obtain a positive variance. This makes precise our earlier remark that most of the p_i should be rather large.

The Poisson Counting Distribution

A good theoretical case can be made for the Poisson density:

$$\pi_n = \frac{\lambda^n e^{-\lambda}}{n!}, \; (n = 0, 1, \ldots) \tag{18}$$

as an appropriate choice for the counting law; Gerber (1979) presents an argument based on a limiting result from the fixed portfolio model, as well as an argument based upon a dynamic portfolio, in which claiming policies are immediately replaced by equivalent, non-claiming policies. (16) then leads to the natural choice

$$\lambda = \Sigma q_j . \tag{19}$$

But from (17) it can immediately be seen that the Poisson assumption, which means $\boldsymbol{V}(\tilde{n}) = \lambda$, leads to *too large* a value of $\boldsymbol{V}(\tilde{n})$ for the second moments (7) and (12) to match. In fact, the collective approximation will now have a variance

$$\boldsymbol{V}(\tilde{y}) = \Sigma q_i v_i + \Sigma q_i (m_i)^2, \tag{20}$$

which is greater than the correct value (7) by the amount $\Sigma q_i^2 m_i^2$.
Another, less critical, problem is that the probability of no claim in the risk approximation:

$$g(0) = \pi_0 = e^{-\lambda} = e^{-\Sigma q_j}, \tag{21}$$

is termwise greater than the true value from (4):

$$g(0) = \prod_{j=1}^{N} p_j . \tag{22}$$

Discussion

In addition to having a good fit between the approximation and the original model, we would also like to have the computation of $g(y)$ via (9), and of related functions, to be efficient; Gerber (1980) describes some of the traditional approximations to the compound Poisson law which have been used by actuaries.
However, a simple recursive scheme for the Poisson case, apparently due to Adelson (1966), has recently been promoted as the most efficient solution to (9). In our notation, it can be shown that:

$$\begin{gathered} g(0) = e^{-\lambda} \\ g(y) = (\lambda / y) \sum_{x=1}^{\min(y,R)} x f(x)\, g(y-x). \qquad (y = 1, 2, \ldots) \end{gathered} \tag{23}$$

This enables exact values of $g(0)$, $g(1)$, $g(2)$, ... to be calculated successively, in a number of steps much less than direct ways of calculating (9). A simple proof of (23), due to Bühlmann and Gerber, can be found in Gerber (1980). Applications can be found in Panjer (1980) and in Held (1980).

More recently, Panjer (1981), has extended the recursive computation of $g(y)$ to a larger class of counting distributions, namely to π_n that are (1) Poisson, (2) Binomial or (3) Negative Binomial (Pascal) (See also Sundt & Jewell (1981) for generalizations).

The Binomial Counting Distribution

From (16) (17), we know that for our problem we want the variance of $\tilde{n}$ to be smaller than the mean; this suggests an improved approximation might result from using a Binomial counting density:

$$\pi_n = \binom{M}{n} \pi^n (1-\pi)^{M-n}, \; (n = 0, 1, \ldots M) \tag{24}$$

with moments

$$\mathbf{E}(\tilde{n}) = \pi M; \; \mathbf{V}(\tilde{n}) = \pi M(1-\pi). \tag{25}$$

For this counting law, Panjer (1981) shows that (23) is replaced by:

$$g(0) = (1-\pi)^M$$

$$(26)$$

$$g(y) = \left(\frac{\pi}{1-\pi}\right) \sum_{x=1}^{\min(y,R)} [(M+1)(x/y)-1] f(x) g(y-x), \; (y = 1, 2, \ldots MR)$$

so that the recursive computation is still more efficient than using (9).
Note especially that we are *not* proposing to set $M = N$, so that both parameters (M, π) are available to match (16) (17). For an exact match of the first two moments, we require that:

$$M = (\Sigma q_i m_i)^2 / (\Sigma q_j^2 m_j^2); \; \pi M = \Sigma q_j. \tag{27}$$

(The reader may easily show that $M \leqq N$.)

However, the Binomial recursive algorithm only works for M integer, so the value obtained above must be rounded up or down, and then π readjusted to provide an exact fit to the first moment. The variance of $\tilde{n}$ and of $\tilde{y}$ will be slightly too large (too small), compared with (16) (17), if M is adjusted upwards (downwards) from the exact value. But this error is in general quite small for moderate values of N.

The Binomial counting distribution also has a good theoretical justification, for if the original portfolio is, in fact, homogeneous, with $q_i = q_0$ and $m_i = m_0$ for all $i = 1, 2, \ldots N$, then we have π_n exactly Binomial, with $M = N$ and $\pi = q_0$. For *small* inhomogeneities, if we set:

$$m_i = m + \mu_i;\; q_i = q_0 + \xi_i;\; (i = 1, 2, \ldots N) \tag{28}$$

where m is defined in (14), and

$$q_0 = \Sigma q_j / N; \tag{29}$$

we find that, to first-order terms in μ_i and ξ_i:

$$\pi \approx q_0 \left[1 + (2/N) \Sigma(\mu_i/m)\right]; \tag{30}$$

$$M \approx N \left[1 - (2/N) \Sigma(\mu_i/m)\right]; \tag{31}$$

that is, only the small inhomogeneities in m_i affect the values of M and π.

If our original portfolio becomes quite large ($N \to \infty$), but the policy characteristics (q_i, m_i) remain comparable, then (27) implies that M is of order N and will thus increase without limit, but that π will remain relatively stable. This means that we do *not* expect that, in the limit, our Binomial counting law will become approximately Poisson (which would require $M \to \infty$, $\pi \to 0$, with $M\pi = \lambda$). The justification of the Poisson law thus requires other limiting conditions.

Associated Functions

As pointed out by Gerber (1980), once a recursive procedure for the density $g(y) = \Pr\{\tilde{y} = y\}$ has been set up, it is a trivial matter to initialize and calculate other associated functions. The functions which seem of most interest are the complementary distribution function:

$$G^c(y) = \Pr\{\tilde{y} > y\} = \sum_{x=y+1}^{\infty} g(x) = 1.0 - \sum_{x=0}^{y} g(x),$$

and the "stop-loss premium":

$$\Pi_{sl}(y) = \boldsymbol{E}\{(\tilde{y} - y)^+\} = \sum_{x=y}^{\infty} G^c(x) = \boldsymbol{E}(\tilde{y}) - \sum_{x=0}^{y-1} G^c(x).$$

Extension to Negative Discrete Values

Adelson's and Panjer's algorithms were developed only for positive $\tilde{w}_i$, which is why the above discussion was limited to the sum of non-negative $\tilde{x}_i$. However, Sundt & Jewell (1981) indicate how arbitrary values, say $\tilde{x}_i$ in the range $[-L, \ldots, 0, \ldots R]$ $(L, R>0)$, can, in principle, be handled for π_n Poisson or Binomial; we develop only the Binomial case.

First of all, (2) is replaced by:

$$f_i(x) = \Pr\{\tilde{x}_i = x \mid \tilde{x}_i \neq 0\}, \tag{32}$$

and (26) is replaced by:

$$g(y) = \left(\frac{\pi}{1-\pi}\right) \sum_{x=A}^{B} [(M+1)(x/y)-1]f(x)g(y-x), \; (-ML \leqq y \leqq MR;\; y \neq 0) \tag{33}$$

where $A = \max(y-MR, -L)$, and $B = \min(y+ML, R)$.

$g(0)$ is no longer calculable explicitely from this form, but both $g(-ML)$ and $g(MR)$ are available from first principles, and (33) can be re-arranged to start the recursion at either end. Starting from the lower end, we obtain:

$$g(y) = \begin{cases} 0 & (y < -ML) \text{ and } (y > MR) \\ [\pi f(-L)]^M & (y = -ML) \\ \left[\dfrac{1}{(y+ML)f(-L)}\right]\left[-(\pi^{-1}-1)(y-L)g(y-L) \right. & \\ \qquad \left. + \displaystyle\sum_{x=1}^{C} [(M+1)x - ML - y]f(x-L)g(y-x)\right] & \text{(otherwise)}, \end{cases} \tag{34}$$

with $C = \min(y+ML, L+R)$.

Of course, if L is very large, then there are obvious problems with the accumulation of round-off error, especially if $f(-L)$ and nearby values are small. One can also imagine multi-pass recursive procedures, or iterative techniques using (33) to resolve these numerical-analytic difficulties.

Further Improvements

One can readily imagine a variety of further improvements to the Binomial compound law to provide a better approximation: for example, since (9) is linear in the (π_n), one could take a linear mixture of several counting distributions, and then mix the results of the corresponding recursively calculated aggregate claim densities; this would enable matching higher moments or other attributes of the true density (4).

One direction which we have examined is to provide a better fit to the true value of $g(0) = \Pi p_i$, which, as previously mentioned, is too large in the Poisson case; the Binomial law, $g(0) = \pi_0 = (1-\pi)^M$, seems to give a better numerical fit, but we cannot guarantee this.
In Sundt & Jewell (1981), it is shown how to modify the Panjer algorithm so that the new counting density (π_n') can take on values:

$$\pi_n' = \begin{cases} \varrho + (1-\varrho)\pi_0 & (n=0) \\ (1-\varrho)\pi_n & (n=1, 2, \ldots) \end{cases} \tag{35}$$

where the π_n are Binomial (π, M). Alternatively, one can continue to use (26), and mix the resulting density in the obvious way with the degenerate density at zero. This *modified Binomial compound law* gives us three degrees of freedom (ϱ, π, M).
Assuming that claim amounts are positive, we can match $g(0)$ by

$$g(0) = \pi_0' = \varrho + (1-\varrho)(1-\pi)^M = \Pi p_i, \tag{36}$$

and (25) becomes

$$\boldsymbol{E}(\tilde{n}) = (1-\varrho)\pi M; \; \boldsymbol{V}(\tilde{n}) = (1-\varrho)\left[\pi M(1-\pi) + \varrho\pi M^2\right]. \tag{37}$$

These must be matched numerically to the true values (22) (16) (17) by iterative numerical methods, which we shall not describe. As before, the integrality of M means that we cannot *exactly* match *both* the second moment and $g(0)$, so that one has to decide which improvement is more important.
We shall see in the example to follow that this modified Binomial provides only a modest improvement over the Binomial, and suggests that further refinements will be of marginal value.

A Numerical Example

To illustrate the effect of the approximation improvement, we use a numerical example due to Gerber (1979), in which there are $N = 31$ policies, and the random values $\tilde{x}_i$ are either 0 or a "face value", c_j, with probability p_i or q_i, respectively, as shown in Table I (the duplication of identical policies is typical). Thus $m_i = c_i$ and $v_i = 0$ $(i = 1, 2, \ldots 31)$.

Table I. Number of Policies with Indicated q_i and c_j

	Face Values c_j				
q_i	1	2	3	4	5
.03	2	3	1	2	–
.04	–	1	2	2	1
.05	–	2	4	2	2
.06	–	2	2	2	1

The exact values of the density $g(y)$, the complementary distribution $G^c(y)$, and the stop-loss premium $\Pi_{sl}(y)$, were obtained by convolving 31 two-point $(0, c_j)$ densities, and are given in the first column of Tables IV, V, and VI. From (6) (7), we find that the first two moments of the original portfolio are:

$$\boldsymbol{E}(\tilde{y}) = 4.49; \quad \boldsymbol{V}(\tilde{y}) = 15.3003$$

and that
$$g(0) = 0.23819.$$

The unnormalized prototypical claim density (13) used in both collective risk approximations is shown in Table II.

Table II. Density of Equivalent Homogeneous Claims

x	1	2	3	4	5
$1.4 f(x)$	.06	.35	.43	.36	.20

The first two moments of this "severity" density are:

$$m = 3.207143; \; v = 1.207092.$$

Thus, from (16) (17), the "counting" density moments for an *exact* fit of a collective risk approximation must be:

$$\boldsymbol{E}(\tilde{n}) = 1.4; \ \boldsymbol{V}(\tilde{n}) = 1.323224.$$

Three approximations were computed using recursions (23) and (26) and the method of (35), giving the numerical matching shown in Table III.

Table III. Value Matching for Numerical Example (differing digits underlined).

	Exact Values	Approximations		
		Poisson	Binomial	Modified Binomial
$\boldsymbol{E}(\tilde{y})$	4.49	4.49	4.49	4.49
$\boldsymbol{V}(\tilde{y})$	15.3003	$1\underline{6.0900}$	$15.3\underline{146}$	15.3003
$\Pr\{\tilde{y}=0\} = g(0)$	0.23819	$0.2\underline{4660}$	$0.23\underline{714}$	$0.238\underline{09}$

In the Poisson approximation, $\lambda = 1.4$ fixed $\boldsymbol{E}(\tilde{y}) = 4.49$ as desired but $\boldsymbol{V}(\tilde{y}) = 16.0900$ and $g(0) = 0.24660$ are significantly too large. Results using the recursion (23) (part of which were also given in Gerber (1979)) are shown in column two of Tables IV, V, and VI.

For the Binomial counting distribution, an exact match of the first two moments would require $M = 25.528480$ and $\pi = 0.0548400$. Rounding up, we select integer $M = 26$, and adjust $\pi = 0.0538462$ to keep $\boldsymbol{E}(\tilde{y}) = 4.49$. $\boldsymbol{V}(\tilde{y}) = 15.3146$ is still significantly close to the exact value of 15.3003, but $g(0) = 0.23714$ is now less than the true value. Note that the range of the Binomial approximation extends, in principle, to $5 \times 26 = 130$, whereas the largest possible total claim sum of the original portfolio is only 97. However, reference to Table V shows that the probability of a claim larger than 40 is already of order 10^{-9}!

For the modified Binomial approximation, we must use (36) (37) to find the parameter values to match the first two moments and $g(0)$; These turn out to be $M = 21.737130$, $\pi = 0.0648672$, and $\varrho = 0.00711084$. Rounding up, we set $M = 22$, and readjust the other values to match the mean and variance, giving finally $\pi = 0.064055$ and $\varrho = 0.00653874$. As can be seen from Table III, the resulting mismatch in $g(0)$ is quite small.

Table IV. Total Sum Densities in Example (differing digits underlined)

	$g(y) = \Pr\{\tilde{y}=y\}$			
	EXACT	APPROXIMATIONS		
Y	RESULT	POISSON	BINOMIAL	MODIFIED BINOMIAL
0	0.23819	0.24660	0.23714	0.23809
1	0.01473	0.01480	0.01504	0.01494
2	0.08773	0.08675	0.08818	0.08762
3	0.11318	0.11122	0.11313	0.11246
4	0.11071	0.11040	0.11256	0.11206
5	0.09633	0.09286	0.09507	0.09492
6	0.06155	0.06101	0.06291	0.06315
7	0.06902	0.06543	0.06732	0.06759
8	0.05482	0.05458	0.05589	0.05613
9	0.04315	0.04132	0.04197	0.04217
10	0.03011	0.03058	0.03071	0.03086
11	0.02353	0.02331	0.02311	0.02321
12	0.01828	0.01834	0.01797	0.01802
13	0.01251	0.01315	0.01265	0.01266
14	0.00871	0.00922	0.00866	0.00865
15	0.00591	0.00650	0.00596	0.00593
16	0.00415	0.00460	0.00411	0.00408
17	0.00272	0.00318	0.00277	0.00273
18	0.00174	0.00212	0.00179	0.00176
19	0.00112	0.00141	0.00115	0.00112
20	0.00071	0.00094	0.00073	0.00071
30	3.09434×10^{-6}	8.63294×10^{-6}	3.98500×10^{-6}	3.51483×10^{-6}
40	3.53514×10^{-9}	36.4155×10^{-9}	7.37055×10^{-9}	5.46425×10^{-9}

Table V. Complementary Distributions in Example (differing digits underlined)

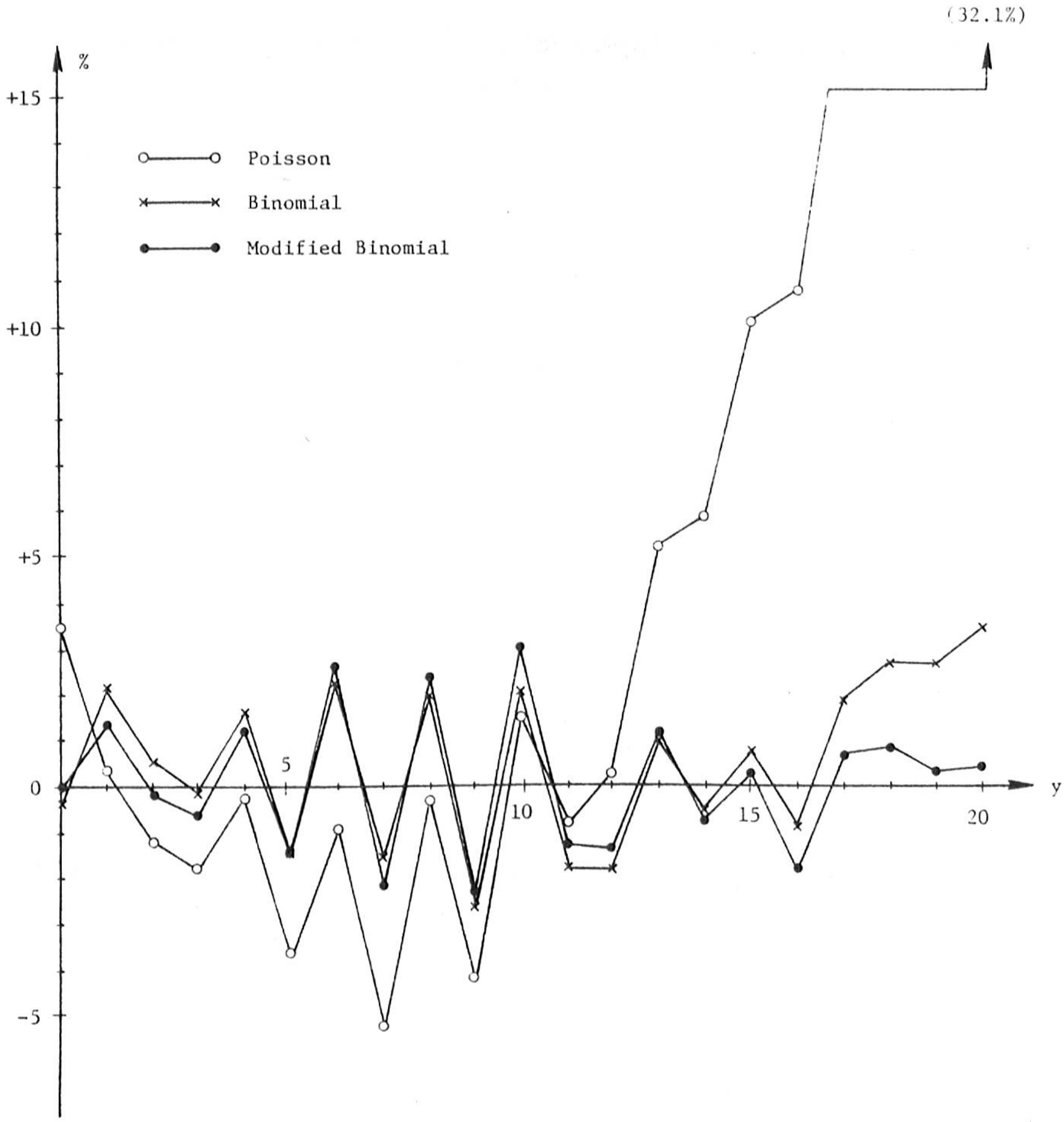
	$G^c(y) = \Pr\{\tilde{y}>y\}$			
	EXACT	APPROXIMATIONS		
y	RESULT	POISSON	BINOMIAL	MODIFIED BINOMIAL
0	0.76181	0.75340	0.76286	0.76191
1	0.74707	0.73861	0.74782	0.74696
2	0.65934	0.65185	0.65964	0.65934
3	0.54615	0.54063	0.54651	0.54688
4	0.43544	0.43023	0.43395	0.43482
5	0.33912	0.33737	0.33888	0.33990
6	0.27757	0.27637	0.27597	0.27675
7	0.20855	0.21094	0.20865	0.20916
8	0.15373	0.15636	0.15276	0.15303
9	0.11058	0.11504	0.11079	0.11086
10	0.08048	0.08446	0.08008	0.08000
11	0.05695	0.06115	0.05696	0.05679
12	0.03866	0.04281	0.03899	0.03877
13	0.02615	0.02966	0.02635	0.02611
14	0.01744	0.02044	0.01769	0.01746
15	0.01153	0.01394	0.01173	0.01153
16	0.00738	0.00934	0.00762	0.00745
17	0.00467	0.00617	0.00485	0.00472
18	0.00292	0.00404	0.00306	0.00296
19	0.00181	0.00263	0.00192	0.00184
20	0.00110	0.00169	0.00118	0.00112
30	3.49840×10^{-6}	12.4621×10^{-6}	4.87524×10^{-6}	4.16710×10^{-6}
40	3.10833×10^{-9}	45.5298×10^{-9}	7.42541×10^{-9}	5.26013×10^{-9}

Table VI. Stop-Loss Premiums in Example (differing digits underlined)

	$\Pi_{sl}(y) = E[(\tilde{y}-y)^+]$			
	EXACT	APPROXIMATIONS		
y	RESULT	POISSON	BINOMIAL	MODIFIED BINOMIAL
0	4.49000	4.49000	4.49000	4.49000
1	3.72819	3.73660	3.72714	3.72809
2	2.98112	2.99799	2.97932	2.98113
3	2.32179	2.34614	2.31968	2.32179
4	1.77563	1.80551	1.77317	1.77491
5	1.34019	1.37527	1.33922	1.34009
6	1.00106	1.03790	1.00034	1.00019
7	0.72350	0.76153	0.72437	0.72345
8	0.51495	0.55059	0.51572	0.51428
9	0.36122	0.39423	0.36296	0.36125
10	0.25064	0.27919	0.25217	0.25039
11	0.17017	0.19472	0.17209	0.17039
12	0.11322	0.13357	0.11513	0.11360
13	0.07456	0.09076	0.07614	0.07483
14	0.04840	0.06110	0.04979	0.04872
15	0.03096	0.04065	0.03210	0.03126
16	0.01943	0.02671	0.02037	0.01973
17	0.01205	0.01737	0.01276	0.01228
18	0.00738	0.01120	0.00791	0.00756
19	0.00446	0.00716	0.00485	0.00460
20	0.00265	0.00453	0.00293	0.00276
30	7.25353×10^{-6}	29.7953×10^{-6}	10.5809×10^{-6}	8.88376×10^{-6}
40	5.72441×10^{-9}	101.020×10^{-9}	14.6686×10^{-9}	10.1485×10^{-9}

Comparison of the different results for the density, $g(y)$, in Table IV shows, as expected, that none of the approximations is a particularly good point estimator; because of the differences in the models, the approximations are forced to fluctuate above and below the exact density. The modified Binomial is generally better than the Binomial, which is generally better than the Poisson, although this is by no means uniformly true.

Figure 1. Percentage Error in Approximations to Density $g(y)$ versus y.

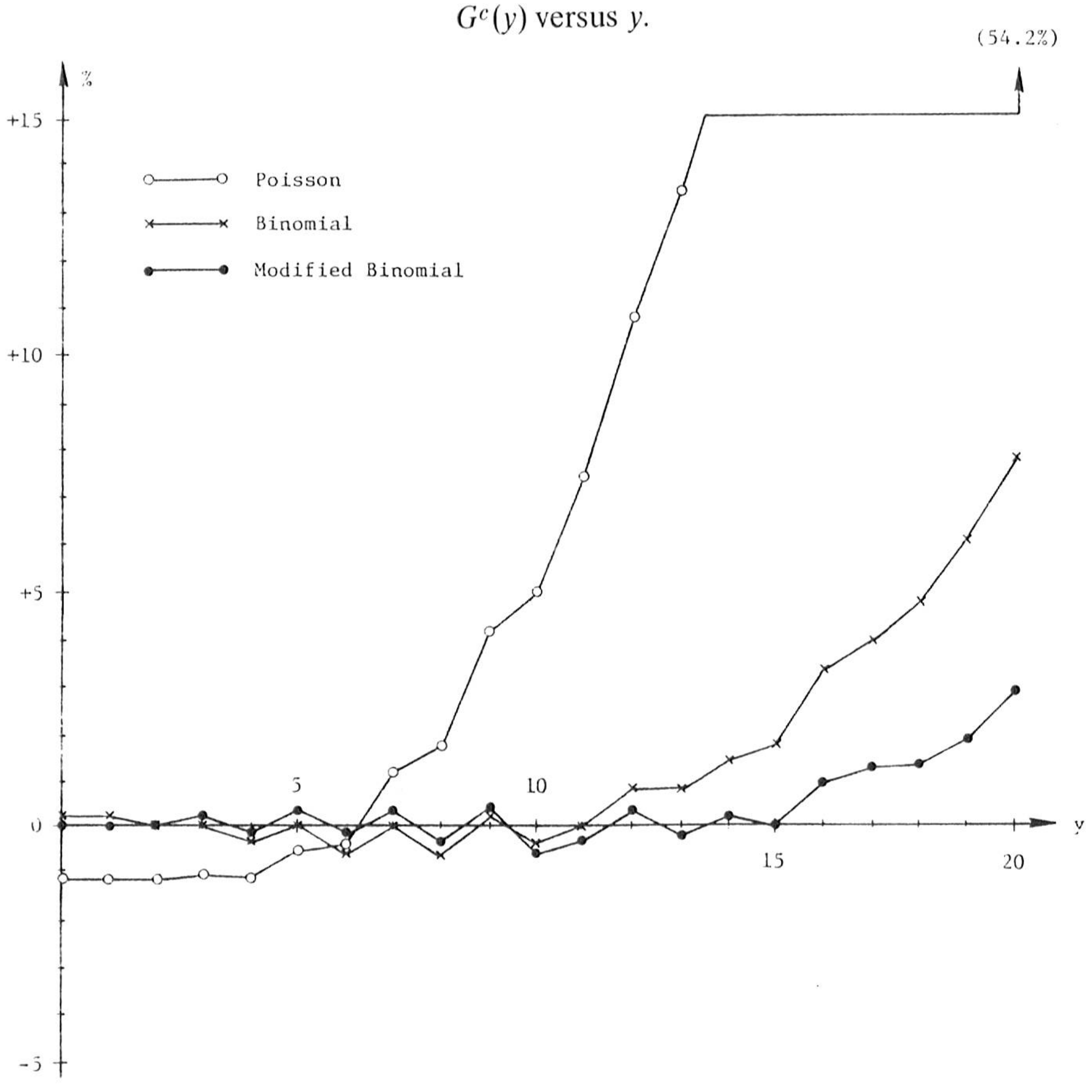


However, when we examine the complementary distributions, $G^c(y)$, in Table V, the approximations become more stable, and the Binomial is always better than the Poisson, except for $y = 6$. The modified Binomial is uniformly best only from $y = 12$ onwards.

The approximations to the stop-loss premiums, $\Pi_{sl}(y)$, in Table VI, are even more stable, and show clearly the value of matching the second moment for this "tail of the tail". The Poisson is always worst, and the modified Binomial always best, except at $y = 6$.

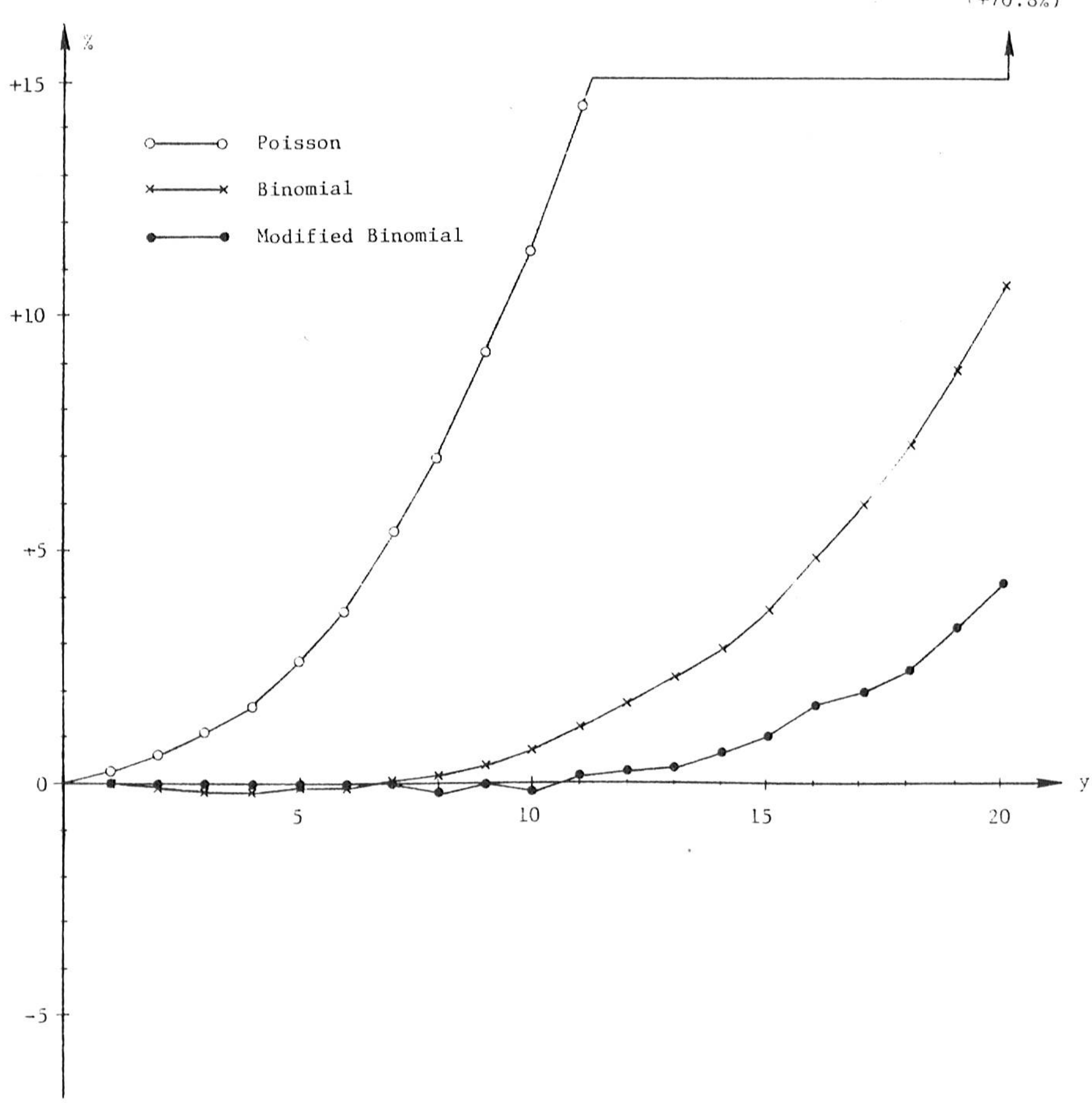
These remarks can be more easily visualized in Figures 1, 2, and 3, which show the *percentage* error in each approximation for the functions of interest.

Figure 2. Percentage Error in Approximations to Complementary Distribution $G^c(y)$ versus y.



In addition to the remarks above, it is of interest to observe the inevitable degradation of all approximations at large values of y. It can be shown theoretically (Bühlmann, et al. 1977) that the Poisson approximation gives too conservative (large) a value for the stop-loss premium for all values of y. Our example suggests all of these approximations are eventually "too dangerous" in the tails. However, it should be remembered that the actual values of the probabilities and of the absolute errors are quite small above $y = 20$.

Figure 3. Percentage Error in Approximations to Stop-Loss Premium, $\Pi_{sl}(y)$ versus y.



Conclusions

Naturally, only limited conclusions can be drawn from a single computational example. However, we believe that the Binomial compound law is a significantly better approximation to the distribution of the original heterogeneous portfolio than the traditional Poisson compound law approximation; furthermore, it can also be computed recursively with little increase in difficulty. There also seems to be evidence that the slight additional work to set up the modified Binomial compound law approximation will be worthwhile if more accurate values of the complementary distribution or the stop-loss premium are desired in the tails.

It remains to be seen whether there are significant differences between these approximations for real risk portfolios, where N and R are both large, and where round-off error accumulation may become important in any recursive method. There have been some claims that other approximation methods or fast Fourier transforms may be competitive under these conditions.

Finally, we must keep in mind the ever-increasing capabilities of digital computers, and the fact that many real portfolio distributions can best be calculated directly.

William S. Jewell
IEOR, 4173 EH
University of California
Berkeley, CA 94720
USA

Bjørn Sundt
Forschungsinstitut für
Mathematik
ETH-Zentrum
8092 Zürich

References

[1] *Adelson, R.M.:* "Compound Poisson Distributions," *Operations Research Quarterly, 17,* pp. 73–75 (1966).

[2] *Bühlmann, H., Gagliardi, B., Gerber, H.U. and Straub, E.:* "Some Inequalities for Stop-Loss Premiums." *ASTIN Bulletin, 9,* pp. 75–83 (1977).

[3] *Gerber, H.U.: An Introduction to Mathematical Risk Theory,* Monograph No. 8, Huebner Foundation, R.D. Irwin, Inc. Homewood, Illinois (1979).

[4] – "On the Numerical Evaluation of the Distribution of Aggregate Claims and its Stop-Loss Premiums," Risk Theory Conference, Mathematics Research Center, Oberwolfach, October 13–17, 1980.

[5] *Held, R.P.:* "Zur numerischen Berechnung von Prämien für die Überschaden-Rückversicherung autonomer Pensionskassen," Diplomarbeit, Höhere Fachprüfung für Pensionsversicherungsexperten, March–September, 1980.

[6] *Panjer, H.H.:* "The Aggregate Claims Distribution and Stop-Loss Reinsurance," Transactions of the Society of Actuaries, pp. 61–73 (1980). (See also Discussion.)

[7] – "Recursive Evaluation of a Family of Compound Distributions." *ASTIN Bulletin, 12,* pp. 22–26 (1981).

[8] *Sundt, B. and Jewell, W.S.:* "Further Results on Recursive Evaluation of Compound Distributions." *ASTIN Bulletin, 12,* pp. 27–39 (1981).

Summary

A traditional actuarial method for the difficult task of finding the exact distribution of a heterogeneous portfolio approximates the distribution with a compound Poisson law with identically distributed risk. This paper shows that a Binomial compound law provides a better match to the second moment of the distribution, thus giving a better approximation, while retaining a simple, recursive algorithm for calculating the distribution. A modified Binomial compound law further refines the approximation, with slight additional work.

Zusammenfassung

Eine traditionelle Methode zur Lösung des schwierigen Problems der Bestimmung der exakten Verteilung eines Gesamtschadens besteht darin, diese Verteilung durch eine zusammengesetzte Poissonverteilung mit gleichverteilten Risiken zu approximieren. Die vorliegende Arbeit zeigt, dass eine zusammengesetzte Binomialverteilung eine bessere Übereinstimmung der zweiten Momente und damit eine bessere Approximation liefert, wobei die Methode zur Berechnung dieser Verteilung auf einem einfachen rekursiven Algorithmus beruht. Eine modifizierte zusammengesetzte Binomialverteilung verfeinert die Approximation mit nur geringem Mehraufwand.

Résumé

Il est de tradition, pour résoudre le problème difficile consistant à déterminer la distribution exacte d'un portefeuille hétérogène, de construire un modèle sur la base de la loi de Poisson composée avec des sinistres distribués identiquement. Le présent article montre qu'une loi binomiale composée fournit une meilleure correspondance au niveau du second moment de la distribution, livrant ainsi une meilleure approximation, en combinaison avec un algorithme récursif simple pour le calcul de la distribution. De plus, une loi binomiale composée modifiée permet d'améliorer encore l'approximation et cela sans grand travail supplémentaire.