TRANSACTIONS OF SOCIETY OF ACTUARIES
1966 VOL. 18 PT. 1 NO. 52

EXPANSION OF PROBABILITY DENSITY FUNCTIONS AS A SUM OF GAMMA DENSITIES WITH APPLICATIONS IN RISK THEORY

NEWTON L. BOWERS, JR.

INTRODUCTION

THE basic idea in this paper follows from the observation that, in calculations involving the distribution of total claims in a risk-theory setting, several authors have found it desirable to approximate a risk-theoretic distribution by use of the incomplete gamma function. In a recent paper, D. K. Bartlett [1] evaluated an excess loss ratio premium by use of a gamma distribution whose first and second moments agreed with those of the risk-theoretic loss distribution. Professor D. A. Jones, in his discussion of the Bartlett paper, pointed out that a translated gamma distribution allowed one to make the approximate distribution agree with the risk-theoretic distribution as far as the first three moments are concerned. The method presented in this paper allows one to make as many moments of the approximate distribution as desired agree with those of the risk-theoretic loss distribution.

In Section I of this paper, the formulas for an expansion of a probability density function as a sum of gamma densities will be developed. The formulas presented will allow one to make the first five moments agree, but the method is easily extended to include more moments. In Section II, the result of Section I is adapted so as to be of use in a risk-theoretic setting. Formulas are developed for the probability distribution function and for the stop-loss or excess loss ratio premium. In Section III, two sample calculations are made: the first will be an example given by Cramér for which exact values are known, and the second will be the Bartlett example. Finally, in Section IV, a peculiarity in the result for the Bartlett example will lead to a criteria for a "best" stop-loss level. Net premiums for two such coverages will then be presented. An appendix to the paper shows another characterization of the type of approximation developed in Section I, and several remarks are made regarding the accuracy of the approximation.

I. METHOD OF EXPANSION

The idea for this expansion is based on the Gram-Charlier series which is described in Mood [6, p. 118]. In the Gram-Charlier series, a density function is expanded in terms of the normal density and its derivatives.

The polynomials which multiply $[1/\sqrt{(2\pi)}] \exp\{-x^2/2\}$ are the Hermite polynomials described in some detail in Jackson [4, p. 176]. They are known to be orthogonal on the entire real axis with respect to the weight function $\exp\{-x^2/2\}$.

Consider the Laguerre polynomials given by

$$L_n^{(a)}(x) = (-1)^n x^{1-a} e^x \frac{d^n}{dx^n}(x^{n+a-1} e^{-x}).$$

The form of the Laguerre polynomials involves a slight change in notation from that usually used and is tied in with the form of the gamma density used in this paper:

$$f(x) = \frac{x^{a-1} e^{-x}}{\Gamma(a)}.$$

One advantage of this form of the gamma density as opposed to

$$g(x) = \frac{x^a e^{-x}}{a!}$$

is that the mean and variance of a random variable with the density $f(x)$ are both a. This change seemed to cut down the length of the formulas in the paper slightly. The first six of these polynomials as defined above are shown below:

$$\begin{aligned}
L_0^{(a)}(x) &= 1\,, \\
L_1^{(a)}(x) &= x - a\,, \\
L_2^{(a)}(x) &= x^2 - 2(a+1)x + (a+1)a\,, \\
L_3^{(a)}(x) &= x^3 - 3(a+2)x^2 + 3(a+2)(a+1)x - (a+2)(a+1)a\,, \\
L_4^{(a)}(x) &= x^4 - 4(a+3)x^3 + 6(a+3)(a+2)x^2 - 4(a+3)(a+2) \\
&\qquad \times (a+1)x + (a+3)(a+2)(a+1)a\,, \\
L_5^{(a)}(x) &= x^5 - 5(a+4)x^4 + 10(a+4)(a+3)x^3 - 10(a+4)(a+3) \\
&\qquad \times (a+2)x^2 + 5(a+4)(a+3)(a+2) \\
&\qquad \times (a+1)x - (a+4)(a+3)(a+2) \\
&\qquad \times (a+1)a\,.
\end{aligned} \tag{1}$$

The pattern of these polynomials is clear at this point. Further, these polynomials are known [4, p. 184] to be orthogonal on the positive real axis with respect to the weight function

$$\frac{x^{a-1} e^{-x}}{\Gamma(a)}.$$

This means that

$$\frac{1}{\Gamma(a)}\int_0^\infty z^{a-1}e^{-z}L_n^{(a)}(z)L_m^{(a)}(z)\,dz=0\,, \qquad \text{if} \qquad m\neq n\,.$$

Furthermore,

$$\frac{1}{\Gamma(a)}\int_0^\infty z^{a-1}e^{-z}[L_n^{(a)}(z)]^2dz=\frac{n!\Gamma(a+n)}{\Gamma(a)}\,.$$

Therefore, if we assume that a given density function $f(x)$ may be written as

$$f(x)=\frac{x^{a-1}e^{-x}}{\Gamma(a)}[A_0L_0^{(a)}(x)+A_1L_1^{(a)}(x)+A_2L_2^{(a)}(x)+\ldots],$$

we may use the above orthogonality conditions to determine the A_n. For

$$\begin{aligned}\int_0^\infty f(z)L_n^{(a)}(z)\,dz&=\int_0^\infty\frac{z^{a-1}e^{-z}}{\Gamma(a)}[A_0L_0^{(a)}(z)+A_1L_1^{(a)}(z)\\&\qquad+\ldots]L_n^{(a)}(z)\,dz\\&=A_n\frac{n!\Gamma(a+n)}{\Gamma(a)}\,,\end{aligned}$$

as all other terms equal zero. Therefore,

$$A_n=\frac{\Gamma(a)}{n!\Gamma(a+n)}\int_0^\infty f(z)L_n^{(a)}(z)\,dz\,.$$

Now let us assume that we have a nonnegative valued random variable Y with a sufficient number of moments. Further, let us define a second random variable X by $X=\beta Y$, where β is chosen so that the mean of X equals the variance of X. We then pick a equal to the common value of the mean and variance of X and evaluate the various constants A_n, where we assume for the purposes of developing these formulas that X has a density function $f(x)$.

$$A_0=\int_0^\infty f(z)L_0^{(a)}(z)\,dz=\int_0^\infty f(z)\,dz=1\,,$$

$$A_1=\frac{1}{a}\int_0^\infty f(z)L_1^{(a)}(z)\,dz=\frac{1}{a}\int_0^\infty f(z)(z-a)\,dz=0\,,$$

since a is the mean of X.

$$\begin{aligned}A_2&=\frac{\Gamma(a)}{2!\Gamma(a+2)}\int_0^\infty f(z)L_2^{(a)}(z)\,dz\,,\\&=\frac{\Gamma(a)}{2!\Gamma(a+2)}\int_0^\infty f(z)[z^2-2(a+1)z+(a+1)a]\,dz=0\,,\end{aligned}$$

from the fact that α is the mean and variance of X. Similarly,

$$A_3=\frac{\Gamma(\alpha)}{3!\Gamma(\alpha+3)}(\mu_3-2\alpha),$$

$$A_4=\frac{\Gamma(\alpha)}{4!\Gamma(\alpha+4)}(\mu_4-12\mu_3-3\alpha^2+18\alpha),$$

and

$$A_5=\frac{\Gamma(\alpha)}{5!\Gamma(\alpha+5)}[\mu_5-20\mu_4-(10\alpha-120)\mu_3+60\alpha^2-144\alpha],$$

where μ_n is the nth moment about the mean of the random variable X. Therefore, the approximate density, a partial sum of the series

$$\frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}[1+A_3L_3^{(\alpha)}(x)+A_4L_4^{(\alpha)}(x)+A_5L_5^{(\alpha)}(x)+\ldots], \quad (2)$$

may be used to evaluate details of the distribution of X and hence corresponding details for Y.

The interesting point of the above formulas is that the A_n may be calculated from one's knowledge of the moments of X even though the exact form of $f(x)$ is unknown. This situation is particularly applicable to calculations on the distribution of the amount of claims paid by an insurance company. The true distribution of claims paid is unknown. However, there are several models of the insurance company, two of which are the collective-risk model and the individual-risk model, and it is possible to evaluate moments of the distribution of claims paid under these two models. Obtaining numerically exact results even of the models themselves is extremely difficult [2, p. 184] except in a few special cases. The method proposed here is a method of approximating the distribution of the claims in the model and thus attempts to do what the Esscher approximation does. The exact distribution of the claims in a model is what is called a risk-theoretic loss distribution in other parts of this paper.

II. FORMULAS FOR THE DISTRIBUTION FUNCTION AND THE STOP-LOSS PREMIUM

We now apply the expansion (2) developed in Section I to obtain formulas for the distribution function and for stop-loss premiums. For simplicity, we shall adopt the following notation:

$$\begin{aligned}
A&=\frac{\mu_3-2\alpha}{3!},\\
B&=\frac{\mu_4-12\mu_3-3\alpha^2+18\alpha}{4!}, \qquad (3)\\
C&=\frac{\mu_5-20\mu_4-(10\alpha-120)\mu_3+60\alpha^2-144\alpha}{5!}.
\end{aligned}$$

Then, if we take four terms in the partial sum, the approximation (2) may be written as

$$\frac{x^{a-1}e^{-x}}{\Gamma(a)}\left[1 + A\frac{\Gamma(a)}{\Gamma(a+3)}L_3^{(a)}(x) + B\frac{\Gamma(a)}{\Gamma(a+4)}L_4^{(a)}(x) + C\frac{\Gamma(a)}{\Gamma(a+5)}L_5^{(a)}(x)\right]. \tag{4}$$

If we now substitute the actual polynomials (1) for $L_n^{(a)}(x)$, we obtain

$$\begin{aligned}&\frac{x^{a-1}e^{-x}}{\Gamma(a)}(1-A+B-C)+\frac{x^{a}e^{-x}}{\Gamma(a+1)}(3A-4B+5C)\\&+\frac{x^{a+1}e^{-x}}{\Gamma(a+2)}(-3A+6B-10C)+\frac{x^{a+2}e^{-x}}{\Gamma(a+3)}(A-4B+10C)\\&+\frac{x^{a+3}e^{-x}}{\Gamma(a+4)}(B-5C)+\frac{x^{a+4}e^{-x}}{\Gamma(a+5)}(C),\end{aligned} \tag{5}$$

which is just the sum of six gamma densities. In general, if it is desired to have the approximate distribution agree with the risk-theoretic distribution for n moments ($n > 2$), the sum of $n + 1$ gamma densities is required. We may thus approximate the distribution function

$$F(x) = \int_0^x f(z)\,dz$$

by

$$\begin{aligned}F(x) \doteq{}& \Gamma(x, a)(1 - A + B - C) + \Gamma(x, a + 1)(3A - 4B + 5C)\\&+ \Gamma(x, a + 2)(-3A + 6B - 10C)\\&+ \Gamma(x, a + 3)(A - 4B + 10C)\\&+ \Gamma(x, a + 4)(B - 5C) + \Gamma(x, a + 5)(C),\end{aligned} \tag{6}$$

where $\Gamma(x, a)$ is the incomplete gamma distribution function given by

$$\Gamma(x, a) = \frac{1}{\Gamma(a)}\int_0^x z^{a-1}e^{-z}dz.$$

Values of this function have been tabulated by Pearson [8] and Salvosa [9].

Integration by parts shows that

$$\Gamma(x, a+1) = \Gamma(x, a) - \frac{x^a e^{-x}}{\Gamma(a+1)}.$$

By use of this, we arrive, after a long calculation, at an alternate expression for $F(x)$. It is

$$\begin{aligned} F(x) \doteq \Gamma(x, a) &- A\left[\frac{x^a e^{-x}}{\Gamma(a+1)} - \frac{2x^{a+1} e^{-x}}{\Gamma(a+2)} + \frac{x^{a+2} e^{-x}}{\Gamma(a+3)}\right] \\ &+ B\left[\frac{x^a e^{-x}}{\Gamma(a+1)} - \frac{3x^{a+1} e^{-x}}{\Gamma(a+2)} + \frac{3x^{a+2} e^{-x}}{\Gamma(a+3)} - \frac{x^{a+3} e^{-x}}{\Gamma(a+4)}\right] \\ &- C\left[\frac{x^a e^{-x}}{\Gamma(a+1)} - \frac{4x^{a+1} e^{-x}}{\Gamma(a+2)} + \frac{6x^{a+2} e^{-x}}{\Gamma(a+3)} - \frac{4x^{a+3} e^{-x}}{\Gamma(a+4)} + \frac{x^{a+4} e^{-x}}{\Gamma(a+5)}\right]. \end{aligned} \tag{7}$$

In this form the close relationship between the proposed approximation method and that implied by the Bartlett method is revealed. Essentially Bartlett worked with the first term on the right-hand side of formula (7).

The problem actually discussed in Bartlett's paper is the evaluation of the stop-loss or excess loss ratio premium, given by

$$\Pi(x) = \int_x^\infty (z - x) f(z)\, dz .$$

Starting again from the approximation (4) for the density and using the integration-by-parts argument above, we obtain after a long calculation that

$$\begin{aligned} \Pi(x) \doteq a[1 - \Gamma(x, a+1)] &- x[1 - \Gamma(x, a)] \\ &- A\left[\frac{x^{a+1} e^{-x}}{\Gamma(a+2)} - \frac{x^{a+2} e^{-x}}{\Gamma(a+3)}\right] \\ &+ B\left[\frac{x^{a+1} e^{-x}}{\Gamma(a+2)} - \frac{2x^{a+2} e^{-x}}{\Gamma(a+3)} + \frac{x^{a+3} e^{-x}}{\Gamma(a+4)}\right] \\ &- C\left[\frac{x^{a+1} e^{-x}}{\Gamma(a+2)} - \frac{3x^{a+2} e^{-x}}{\Gamma(a+3)} + \frac{3x^{a+3} e^{-x}}{\Gamma(a+4)} - \frac{x^{a+4} e^{-x}}{\Gamma(a+5)}\right]. \end{aligned} \tag{8}$$

The first two terms, $a[1 - \Gamma(x, a + 1)] - x[1 - \Gamma(x, a)]$, are those used by Bartlett. Another convenient form for these two terms is

$$\frac{a x^a e^{-x}}{\Gamma(a+1)} - (x - a)[1 - \Gamma(x, a)]. \tag{9}$$

III. TWO EXAMPLES

In the following two examples, we shall use the above approximation method to evaluate details of certain risk-theoretic distributions. In both cases the model is from collective-risk theory, and we shall assume that the number of claims is given by a Poisson distribution. In other words, given that the expected number of claims equals t, the probability that n claims occur equals $e^{-t}t^n/n!$. The assumptions on a population which lead to the Poisson distribution for the number of claims are given by Kahn [5, p. 414]. Further, let $\hat{\mu}_n$ be the nth moment about the origin of the distribution of the claim amount, given that a claim has occurred. Then, the moments about the mean, $\hat{\mu}_1 t$, of the distribution of total claims, given that the expected number of claims is t, are as follows:

$$\begin{aligned} \mu_2 &= \hat{\mu}_2 t\,, \\ \mu_3 &= \hat{\mu}_3 t\,, \\ \mu_4 &= \hat{\mu}_4 t + 3\hat{\mu}_2^2 t^2\,, \\ \mu_5 &= \hat{\mu}_5 t + 10\hat{\mu}_2\hat{\mu}_3 t^2\,. \end{aligned} \tag{10}$$

A derivation of the first four moments is given by Bartlett [1, p. 451], and the fifth moment may be derived in the same manner.

TABLE 1

VALUES OF $10^5\,F(y)$

y	x	Exact Value	2d Moment Gamma	3d Moment Gamma	4th Moment Gamma	5th Moment Gamma	Edgeworth Expansion	Esscher Approximation
0	0	0	0	0	0	0	−64	0
4	2	342	110	144	183	216	264	341
8	4	6,039	5,110	5,683	6,086	6,255	6,165	6,033
12	6	25,385	25,589	26,048	25,861	25,518	25,375	25,370
16	8	53,540	54,687	54,067	53,362	53,166	53,526	53,526
20	10	77,387	77,990	77,146	77,032	77,324	77,373	77,376
24	12	91,172	91,054	90,763	91,120	91,338	91,241	91,168
28	14	97,150	96,839	96,974	97,263	97,234	97,134	97,150
32	16	99,218	99,000	99,231	99,306	99,204	99,160	99,218
36	18	99,814	99,711	99,878	99,840	99,782	99,991	99,814
40	20	99,961	99,922	100,009	99,955	99,944	100,010	99,960

A. *An Example from Cramér* [3]

Let the distribution function for the amount of the claim, given that a claim has occurred, be $1 - e^{-y}$. It is easy to verify that $\hat{\mu}_n = n!$. Further assume that the expected number of claims, t, is 16. Then the moments of the distribution of total claim amount, Y, may be calculated by use of formula (10). They are found to be

$$\mu = 16,\ \mu_2 = 32,\ \mu_3 = 96,\ \mu_4 = 3{,}456,\ \mu_5 = 32{,}640\,.$$

The scaling necessary to make the mean equal to variance is one-half. After this scaling the moments of X are

$$\mu = 8,\ \mu_2 = 8,\ \mu_3 = 12,\ \mu_4 = 216,\ \mu_5 = 1{,}020\,.$$

Applying formula (3), we see the constants to be used in the approximation are $a = 8$, $A = -\frac{2}{3}$, $B = 1$, $C = -1.1$.

Table 1 shows values of $F(y)$ for various values of y. The "exact value" figures are adapted from some in Cramér [3, p. 42], as are those for the two approximation methods. One is a four-term Edgeworth expansion based on the normal distribution and its derivatives, and the other is an Esscher approximation using as the approximating distribution another four-term Edgeworth expansion. Both of these methods are described in some detail by Cramér [3, pp. 31–40]. The remaining figures give the results of the approximation method that we are considering correct to 2, 3, 4, and 5 moments.

If we compare the results, we see the gamma approximations are not particularly good near the mean, 8, of the scaled variable X. In fact, the correction term corresponding to the fifth-moment correction, the C term, is still quite large near the mean. However, the gamma approximations appear superior to the Edgeworth expansion near the tails of the distribution. The Esscher approximation is much more accurate but also much more difficult to apply.

B. *An Example from Bartlett*

This example from Bartlett [1] is for a health insurance coverage where the claim amount is proportional to the number of days of disability. From a table in the Bartlett paper, one can evaluate the moments of the distribution of the number of days of disability.

$$\hat{\mu}_1 = 31.35211, \quad \hat{\mu}_2 = 1{,}861.705, \quad \hat{\mu}_3 = 139{,}531.08,$$
$$\hat{\mu}_4 = 11{,}453{,}147.7, \quad \hat{\mu}_5 = 979{,}479{,}188\,.$$

The distribution of the number of claims is assumed to be Poisson, with the expected number of claims equal to 14.63. Application of the formulas (10), therefore, gives us the following moments of the total claim amount Y measured in days.

$$\mu = 458.68, \quad \mu_2 = 27{,}237, \quad \mu_3 = 2{,}041{,}300,$$
$$\mu_4 = 2{,}393{,}100{,}000, \quad \mu_5 = 570{,}320{,}000{,}000\,.$$

The scaling necessary to make the mean equal the variance is $X = Y/59.381$. After making this adjustment, the moments of X are

$$\mu = 7.7244, \quad \mu_2 = 7.7244, \quad \mu_3 = 9.7491,$$
$$\mu_4 = 192.4737, \quad \mu_5 = 772.4700\,,$$

and the constants of the approximation are

$$\alpha = 7.7244, \quad A = -0.950, \quad B = 1.480, \quad C = -1.604\,.$$

To calculate the stop-loss premium associated with 120 per cent of the pure-risk premium, we evaluate formula (8) with $\alpha = 7.7244$ and $x =$

$1.2\alpha = 9.2693$. The accompanying tabulation gives the stop-loss premium correct to various numbers of moments.

Number of Moments	Stop-Loss Premium $\Pi(x)$	Stop-Loss Premium Scaled to Y
2............	0.55429	32.91
3............	.56163	33.35
4............	.54111	32.13
5............	0.53632	31.85

The figure 32.91 compares with the 32.24 figure which Bartlett obtained by linearly interpolating between figures corresponding to $\alpha = 5$ and $\alpha = 10$ to get the figure for $\alpha = 7.7244$. The figures presented here are based on the tables by Salvosa [9], where in effect the values of α are closer together.

The most interesting feature of this example is the fact that improving the approximation by making calculations correct to the third moment increases the stop-loss premium. In this particular example, the third moment of the gamma distribution fitted by use of the first two moments of the risk-theoretic distribution is $2\alpha = 15.4488$, and the third-moment correction is made to reduce this third moment to 9.7491. In other words, the adjustment for the third moment is made to reduce the skewness of the approximation, and making this change still increases the stop-loss premium. An answer to this apparent paradox is made in Section IV.

IV. SOME "BEST" STOP-LOSS REINSURANCE AGREEMENTS

It was noted in Section III that, in one case at least, an adjustment to reduce skewness increases a stop-loss premium. To see that this is not a peculiarity of the approximation method used, assume that the density function for claims is the gamma density $[1/\Gamma(\alpha)]x^{\alpha-1}e^{-x}$, which has mean and variance equal to α. The stop-loss premium for claims in excess of the mean is given by formula (9) to be $\alpha^{\alpha+1}e^{-\alpha}/\Gamma(\alpha+1)$. For the normal density with mean α and variance α, the stop-loss premium for claims in excess of the mean is $\sqrt{(\alpha/2\pi)}$. But it is known [7, p. 528] that $\Gamma(\alpha+1) = \alpha! > (\alpha/e)^{\alpha}\sqrt{(2\pi\alpha)}$ for $\alpha = 1, 2, 3, \ldots$. Therefore, the stop-loss premium is greater under the normal density (unskewed) than it is under the gamma density, which is skewed positively.

If we examine the third-moment correction in the formula (8) for the stop-loss premium, we notice that this term may be written as

$$-A\frac{x^{\alpha+1}e^{-x}}{\Gamma(\alpha+2)}\left(1-\frac{x}{\alpha+2}\right).$$

Thus for $x < \alpha + 2$, the stop-loss premium is increased when $-A = (2\alpha - \mu_3)/3!$ is positive, that is, when the third moment of X is less than 2α. However, if the stop-loss level is chosen to be $x = \alpha + 2$, then the third-moment correction is equal to zero. In that case, the fourth-moment correction in the formula (8) for the stop-loss premium would be

$$+B\frac{x^{\alpha+1}e^{-x}}{\Gamma(\alpha+2)}\left[1-\frac{2x}{\alpha+2}+\frac{x^2}{(\alpha+2)(\alpha+3)}\right],$$

which, when $x = \alpha + 2$ is substituted, becomes

$$-B\frac{(\alpha+2)^{\alpha+2}e^{-(\alpha+2)}}{\Gamma(\alpha+4)}.$$

In both cases examined in Section III, this quantity was negative. Similarly, the fifth-moment-correction term at $x = \alpha + 2$ in formula (8) may be reduced to

$$+4C\frac{(\alpha+2)^{\alpha+2}e^{-(\alpha+2)}}{\Gamma(\alpha+5)},$$

which was also negative in both examples. This suggests then that we should use the stop-loss level of $x = \alpha + 2$ and base the premium on only the first two moments. There would be no third-moment correction, and our two examples suggest the fourth- and fifth-moment corrections are usually negative.

Now let us assume that we are studying some group or small insurance company and have set up an appropriate collective-risk or individual-risk model for it. Let us assume that μ is the expected amount of claims of this group under our model and that σ is the standard deviation of the amount of claims. Then σ^2 is the variance, and the required scaling is found by setting $\beta\mu = \beta^2\sigma^2$, so that $\beta = \mu/\sigma^2$. Therefore, $\alpha = \beta\mu = \mu^2/\sigma^2$, so that the stop-loss level of $\alpha + 2 = \mu^2/\sigma^2 + 2$ for the scaled variable. Thus, the stop-loss level for the original unscaled variable is $\mu + 2\sigma^2/\mu$. The stop-loss premium for the scaled variable is by formula (8) equal to

$$\Pi(\alpha+2)=\alpha[1-\Gamma(\alpha+2,\,\alpha+1)]-(\alpha+2)[1-\Gamma(\alpha+2,\,\alpha)],$$

which is a function of α alone. But α is equal to μ^2/σ^2, so $\Pi(\alpha + 2)/\alpha$ may be looked on as a function of σ/μ. Since $100\,\Pi(\alpha + 2)/\alpha$ is the stop-loss premium as percentage of the expected claims of the scaled variable and is equal to the same ratio for the original unscaled variable, we may tabulate the stop-loss premium for this coverage in the quite simple manner shown in Table 2.

In what sense is this a "best" stop-loss reinsurance agreement? Primarily because only two moments, the mean and variance of expected total claims, must be calculated for any group or reinsured company to calculate the stop-loss premium. Wide variation in the skewness of the distribution of total claims makes no difference in the stop-loss premium calculated by an approximation correct through three moments. Whether such a three-moment approximation is generally conservative is problematical—it was conservative in the two examples in Section III.

This stop-loss reinsurance scheme was designed from considerations arising from the expansion described in Section I. If a Gram-Charlier expansion is made and the development of Sections I and II followed through, it may be shown that the stop-loss premium has a zero third-moment correction term only if the stop-loss level is the expected claims.

TABLE 2

PREMIUM FOR STOP-LOSS COVERAGE
EXCESS CLAIMS OVER $\mu + 2\sigma^2/\mu$
PREMIUM AS PERCENTAGE OF μ

σ/μ	Premium as Percentage of μ	σ/μ	Premium as Percentage of μ
0.00	0.000%	0.35	5.746%
.05	1.758	0.40	5.832
.10	3.092	0.45	5.852
.15	4.076	0.50	5.825
.20	4.777	0.55	5.768
.25	5.257	0.7071	5.494
0.30	5.565	1.00	4.979

For this case, the stop-loss premium was found to be $\sigma/\sqrt{(2\pi)}$, as indicated at the beginning of this section.

The author would like to take this opportunity to thank Professors C. J. Nesbitt and D. A. Jones, who have read various drafts of this paper and have made many helpful suggestions.

APPENDIX

It is known [4, p. 27] that, in an expansion of a suitable function in a Fourier series, the resulting partial sum at any stage is a trigonometric polynomial which minimizes a certain integral. A similar statement can be made about the expansion developed in Section I.

Assume a is given. We then form the integral

$$G=\Gamma(a)\int_0^{\infty}\frac{1}{z^{a-1}e^{-z}}\left[f(z)-\frac{z^{a-1}e^{-z}}{\Gamma(a)}\sum_{k=0}^{n}B_k z^k\right]^2 dz.$$

G measures the error between $f(x)$ and an approximating density, which is a gamma density times a polynomial. G is formed by squaring this error, multiplying by a weight function $\Gamma(a)/x^{a-1}e^{-x}$ for each x, and then integrating. For $a > 1$, this has the effect of strongly weighting the errors at both tails of the distribution. For $a \leq 1$, the weight function has small effect near $x = 0$ but strongly weights error for large values of x. An attempt is then made to select the B_k so that G is minimized. It proves convenient in this calculation to change the polynomial

$$\sum_{k=0}^{n} B_k x^k$$

into the form

$$\sum_{k=0}^{n} A_k L_k^{(a)}(x),$$

which is again a polynomial of degree at most equal to n. To effect the minimization, we take the partial derivatives $\partial G/\partial A_k$ and set them equal to zero. But $\partial G/\partial A_k = 0$ implies

$$-2\int_0^\infty f(z) L_k^{(a)}(z)\, dz + 2A_k \int_0^\infty \frac{z^{a-1}e^{-z}}{\Gamma(a)} [L_k^{(a)}(z)]^2 dz = 0,$$

where again use is made of the orthogonality of the Laguerre polynomials relative to the weight function $x^{a-1}e^{-x}/\Gamma(a)$. This shows that A_k should be chosen so that

$$A_k = \frac{\int_0^\infty f(z) L_k^{(a)}(z)\, dz}{\int_0^\infty \frac{z^{a-1}e^{-z}}{\Gamma(a)} [L_k^{(a)}(z)]^2 dz} = \frac{\Gamma(a)}{k!\Gamma(a+k)} \int_0^\infty f(z) L_k^{(a)}(z)\, dz.$$

In the sense of the error integral G, adding an additional term improves the approximation as A_{n+1} correctly chosen gives a lower value to the integral than the choice of $A_{n+1} = 0$. Unfortunately, no estimates are known to the author as to bounds on the error in the approximation using this criteria or some other. However, the form of the weight function, low near the mean, suggests that the approximation is less accurate in that range, and the example by Cramér studied in Section III, A, bears this out.

BIBLIOGRAPHY

1. Bartlett, D. K. "Excess Ratio Distribution in Risk Theory," *TSA*, XVII (1965), 435.
2. Bohman, H., and Esscher, F. "Studies in Risk Theory with Numerical Illustrations concerning Distribution Functions and Stop-Loss Premiums," *Skandinavisk Aktuarietidskrift*, XLVI (1963), 173.
3. Cramér, H. "Collective Risk Theory," Jubilee Vol. of *Forsakingsaktiebolaget Skandia*, 1955.

4. JACKSON, D. *Fourier Series and Orthogonal Polynomials.* Menasha, Wis.: Mathematical Association of America, 1941.
5. KAHN, P. M. "An Introduction to Collective Risk Theory and Its Applications to Stop-Loss Reinsurance," *TSA*, XIV (1962), 400.
6. MOOD, A. M. *Introduction to the Theory of Statistics*, 1st ed. New York: McGraw-Hill Book Co., Inc., 1950.
7. OLMSTED, J. M. H. *Real Variables.* New York: Appleton-Century-Crofts, Inc., 1956.
8. PEARSON, K. *Tables of the Incomplete Gamma Function.* London, 1922.
9. SALVOSA, L. R. "Tables of Pearson's Type III Function," *Annals of Mathematical Statistics*, I (1930), 191.

DISCUSSION OF PRECEDING PAPER

PAUL MARKHAM KAHN:

Mr. Bowers' paper represents a very fine contribution to the analysis of claims distributions and the approximations to them using standard functions. Recently this subject has received considerable attention, a notable example of which is the Bohman-Esscher report based on the work of a Swedish committee of actuaries.

One stimulus to the use of the gamma distribution in these investigations is its relative simplicity compared with the Esscher approximation, which, though complicated, gives good results, as shown in Table 1 of Mr. Bowers' paper. The purpose of this discussion is to call attention to a very recent paper by two Finnish actuaries ("Approximations of the Generalised Poisson Function," by Lauri Kauppi and Pertti Ojantakanen), presented to the 1966 ASTIN Colloquium. Kauppi and Ojantakanen give a simple formula for the amount of claims corresponding to a given probability level, and this formula produces results that agree very well with the Esscher approximation.

One note of caution, however, should be made concerning any of these approximations. It seems unreasonable to look for accuracy beyond a few significant figures, especially when calculating levels of security.

WILLIAM D. BERG:

I wish to congratulate Mr. Bowers on his well-written paper. The student of mathematical statistics will find it a pleasure to read, especially as an introduction to the mathematical aspects of collective risk theory. His presentation of the sum of the gamma functions is elegant and illuminating.

Because of the central importance of the expression (9), the insight gained from what I shall call a quadrature interpretation is useful. By writing $1/\Gamma(\alpha)$ for $\alpha/\Gamma(\alpha+1)$, the first term, which is the first approximation to the stop-loss net premium, is more readily seen as the area of the rectangle whose length is x and whose height is the ordinate of the density function at x (x is the scaled-down claim amount). The next approximation reduces this by an amount represented by the area of the rectangle whose length is the excess of x over its expected value α and whose height is the fraction represented by the area of the tail of the p.d.f. beyond the value x. The quadrature interpretation helps one to visualize the effect of a variation in x on the stop-loss net premium.

I trust that the author will catch the typographical error in the galley proof that uses lower case *y*'s for capital *Y*'s in the heading of Table 1. The same slip with x is not material.

B. GEORGE ISEN:

Actually the process of fitting a smooth curve to a sample does not add anything to our information about $f(x)$ that is not contained in the sample. The fitted curve may, in fact, give one an entirely misleading impression of the real density function.[1]

It has been interesting and enlightening to follow the recent incursions made into the hazy never-never land of stop-loss reinsurance. This problem area, and the methods which have been proposed to establish a "science-proves" solution, are related to many other important problems to the actuary, including among them the establishment of contingency reserves, the chance of ruin, the establishment of retention limits for life insurance policies, and certain problems related to group insurance policies and their experience rating.

Mr. Herbert Feay[2] has certainly made the members of the Society indebted to him "for his courage, talent, and industry"[3] in this field. After Mr. Feay thoroughly classified various approaches to reinsurance, its purposes, and some general considerations, he proceeded into the misty bog in the section entitled "Determination of Premiums." This section of the paper evoked many other suggested pathways through the murk and elicited much erudite, though critical, discussion (which, if not critical, is probably not constructive). I hesitate to add my discussion of Mr. Bowers' ably written and informative paper to the long list of writings stimulated by Mr. Feay's paper, including those most important papers by Dr. Paul Markham Kahn[4] and Mr. Dwight K. Bartlett.[5]

However, the line of papers seems to be bearing down on one basic problem: What is the true underlying nature of the risk, how may it best be expressed mathematically, and how may this "mathematic" be applied in day-to-day work? (This is, in reality, only one basic problem, analyzed into its fundamental parts, much as the atom, which is the basis of all matter, may be analyzed into its fundamental parts, etc.) Mr. Feay

[1] Mood, *Introduction to the Theory of Statistics* (New York: McGraw-Hill Book Co., Inc., 1950), p. 120.

[2] Herbert L. Feay, "Introduction to Nonproportional Reinsurance," *TSA*, XII, 22.

[3] Irving Rosenthal, in his discussion of Feay's paper, *TSA*, XII, 54.

[4] "An Introduction to Collective Risk Theory and Its Application to Stop-Loss Reinsurance," *TSA*, XIV, 400.

[5] "Excess Ratio Distribution in Risk Theory," *TSA*, XVII, 435.

seemed interested in establishing the basic, simple mathematical framework of his subject[6] and then, justifying the use of the normal distribution, proceeded to calculate some illustrative premiums. As I intimated above, the discussions then went on to challenge the use of the normal distribution as the basic underlying function describing the true density of the claim distribution. It is not necessary to mention each discussant separately, but it may be of help to partially and hastily (and probably badly) outline the alternative roads to Oz (all varying in their shades of yellow, but all leading, brick by brick, through that country, strange, uncharted, and inhabited only by tin woodmen and straw scarecrows).

The methods which have been suggested include the direct application of collective risk theory (with the use of the Esscher approximation), Monte Carlo techniques, the substitution of the Poisson formula for the normal distribution, and such, thereby stamping Mr. Feay as misguided in the shade of the bricks on his road.

To save the day, Dr. Kahn[7] gave us a summary of the weapons developed and used by our European brethren (seriously, his and Mr. Feay's papers led this poor actuary to the sources so valuably accumulated in the rich bibliographies of their papers), and again Mr. Feay's normal distribution came under attack as being abnormal. Dr. Kahn's pathway is tortuous and narrow, and only the most adroit at compound-Poisson processes may follow in safety.

Mr. Bartlett[8] ventured into the mire, indicating the importance of finding a path and an easy one—like a gamma density function[9] of the type found in chapter vi of Mood.[10] Fitting the first two moments of the actual distribution of claim amounts to those moments of the density function would of necessity determine the value of the parameters of the function and excess ratio probabilities, and, hence, stop-loss premiums could be determined. Mr. Bartlett has not let us travel easily, but, as with the previous scholarly papers of Feay and Kahn, his introduces new concepts which can be evaluated with only the closest attention. The road seems to be narrowing. But, again, out of the discussions can be heard the pleas for

[6] $$_{R}P^{H}_{L}=\int_{t=L}^{t=H}(t-L)f(t)\,dt\times\frac{R}{M}.$$

[7] *Op. cit.*

[8] *Op. cit.*

[9] $$f(x)=\begin{cases}\dfrac{1}{\alpha!\beta^{\alpha+1}}x^{\alpha}e^{-x/\beta} & x>0\\ 0 & x<0\end{cases}.$$

[10] *Op. cit.*, p. 112.

the use of the logarithmico-normal distribution, the Pareto distribution, the T-gamma (with three parameters, and hence the use of three moments) for use as the "secondary" distribution, and weakly, but determinedly, the path opened by a spin of the roulette wheel is again offered for consideration (Monte Carlo and his random numbers). It seems to have been settled that the Poisson distribution leads to the right distribution of the number of claims (except for Mr. Jackson's minor dissent in his discussion of Mr. Bartlett's paper).[11]

Mr. Bowers now invites us to follow along yet another path. He describes very accurately and lucidly the land to which he leads us. He takes advantage of Mr. Bartlett's lead of the use of the gamma function for the density function in question but tends to increase our confidence by invoking the use of more than two or three moments, this being possible through the use of an expansion similar to the Gram-Charlier series, except that the Laguerre polynomials are used as weight functions (this making the solutions neat by their being orthogonal polynomials) in the place of the Hermite polynomials used in the Gram-Charlier series. Mr. Bowers must be congratulated for his thorough analysis and his offer of yet another path to the Emerald City.

In Table 1 of Mr. Bowers' paper a comparison is made of various approximations to an exact value resulting from an explicit function. The Esscher approximation seems to yield the closest values to the exact values, while fifth-moment gamma is certainly not as reliable an estimate to the values in this table. However, it is possible that the fifth-moment gamma function may be justified on other grounds, though it seems that the mathematical work involved in applying the fifth-moment gamma function is approaching that needed to complete the approximation by means of Esscher functions.

As I stand at the edge of this land fraught with dangers, preparing to reach the goal—that city of which only the Wizard of Oz himself knows the location—I see before me the beckoning of many paths: all are yellow and all are brick, but I am at a loss, for my instruction to follow the yellow brick road has guided me to indecision.

It seems that the literature is clear as to what price must be paid when a road is to be selected.[12] In life insurance mathematics, the secret pathway was sought by men of the stature of Gompertz and Makeham. Today, no actuary pretends that these mathematical functions for the force of mortality are anything but techniques of graduation. In the graduation of our most recent mortality table for use in valuation, these were abandoned

[11] Paul H. Jackson, *TSA*, XII, 459.

[12] Mood, *op. cit.*, p. 120.

for, as we all know, the use of Jenkins' fifth-difference modified osculatory interpolation formula for most of the range of ages in the table. At the upper ages, where the data are scarce, artificial methods were used to close off the table.

Mr. Paul Jackson, in his discussion of Kahn's paper,[13] illustrates the possibility of a growing and confusing number of roads to the goal. But he also brings forward a point which should be enlarged upon. The fitting of these many frequency functions depends on the data involved. Of necessity the data are most plentiful at or near the mean of these distributions of actual claim amounts, but there is a paucity of data at the extreme or tail, and here precisely is that magic city that we are trying to reach—the underlying, basic, inherent density function at the tail of the distribution. The problem of evaluating the "true" tail by comparison of a density to the actual data has limited reliability.

The Monte Carlo technique has been advanced, but this technique still relies on a basic underlying probability being associated with various claim amounts, and the analysis of hypothetical samples exposed for a large number of hypothetical time periods, with the occurrence of large amounts being estimated therefrom and probably being small in number at any rate (that is, duplicating the paucity of data in the tail). Because of my limited knowledge of this area, it seems to me that to apply the Monte Carlo technique to, say, a group of lives, requires that ages, probabilities conforming to those ages, and amounts conforming to those lives must be selected and then the results of the simulation analyzed. Here the road is not only a different shade of yellow but has many side lanes in sight even before we enter upon it.

The problem, which seems to be losing its disguise, is not what the underlying function is but rather what is a good graduation formula for the data observed, so that the stop-loss premiums (or other related functions) calculated therefrom may be adequate, equitable, consistent, and competitive. The more parameters added to a graduation formula, the closer we expect the curve to fit to the data. Hence, adding moments and using the "method of moments" (if you will) may give a better fit to the observed data, but does it predict the nature of the tail any better when the data there are so sparse? (Of course, if the basic function does not describe the data closely, then added parameters may cause greater deviation from the observed, this being true where parameters are contained in the terms of a series which itself diverges or converges very slowly.)

Before turning to a conclusion, I would like to review the assumption underlying the Poisson distribution as being representative of the prob-

[13] *TSA*, XIV, pp. 443–44.

ability of the number of claims incurred. Using standard actuarial notation, the probability that (x) dies (or becomes hospitalized or disabled) before age $x + h$ is ${}_hq_x$. Also, let ${}_hp_x = 1 - {}_hq_x$. Then the probability that (x) survives to $x + kh$ and dies before age $x + (k + 1)h$ is ${}_{kh}p_x \cdot {}_hq_{x+kh}$. If it is assumed that $0 \leq h \leq 1$ and k is an integer such that $0 \leq kh < 1$, then ${}_{kh}p_x \cdot {}_hq_{x+kh}$ represents the probability that (x) dies before $x + 1$, and exactly between $x + kh$ and $x + (k + 1)h$. Usually an assumption of uniform distribution of deaths is made. This results in the following:

$$
\begin{aligned}
{}_{kh}p_x \cdot {}_hq_{x+kh} &= \frac{l_{x+kh}}{l_x} \cdot \frac{l_{x+kh} - l_{x+(k+1)h}}{l_{x+kh}} \\
&\doteq \frac{khl_{x+1} + (1 - kh)l_x - (1 - kh - h)l_x - (k + 1)hl_{x+1}}{l_x} \\
&= \frac{h(l_x - l_{x+1})}{l_x} \\
&= h \cdot q_x .
\end{aligned}
$$

Hence, under this assumption the probability that (x) dies in any small part, h, of the year of age is the same as for death in any other part of duration h in that same year. However, the probability of death in a small part of the year, h, if (x) has lived part of the year, increases throughout the year, and this without regard to the size of h. That is:

$$
{}_hq_{x+kh} < {}_hq_{x+(k+1)h} ,
$$

if

$$
h \cdot q_x < {}_{kh}q_{x+(k+1)h} ,
$$

$$
h \cdot q_x < \frac{h \cdot q_x}{p_{x+kh}} ,
$$

or

$$
p_{x+kh} < 1 ,
$$

which it is in ordinary meaning of probability.

Hence, the assumption that the probability of a claim in a short period of time is independent of what happened prior to that time (at least respecting life insurance) is not true, even using the assumption of uniform distribution of claims. (It is obvious where the UDC is not used, because μ_x, the force of mortality, is not considered constant with respect to time in any case respecting life contingencies of mortality or morbidity.) However, where the probability of occurrence in that period is very small and the number exposed to the hazard very large, the Poisson is an excellent approximation to the true probability related to the number of claims. A small number exposed, as in small groups, challenges the accuracy of using the Poisson distribution. I regret that I have not carried forward these thoughts to the point of calculations and comparisons due to lack of time. Mr. Jackson had touched on this problem from another point of view[14] but had not referred to the fact that claims have to be not only independent of each other but also of the time of their occurrence, in order for the Poisson distribution to be exactly representative.

I have digressed sufficiently from making a decision regarding the path to take. In attempting to choose wisely, I have re-examined many basic texts in statistics. Wolfenden's text[15] directed to the problems of actuaries in curve fitting, graduation, and statistics, indicates on page 64 that variations in the Gram-Charlier series are available for use with data that are very skewed. The normal functions may be replaced by the Poisson exponential as the generating function giving rise to the Poisson-Charlier series. Also, the Gram-Charlier series may be used with the logarithmico-normal as the generating function.[16]

In conclusion then, Dorothy was much more successful in reaching the Emerald City than I, for she had only one yellow road to choose. I am still looking for the road with the right shade of yellow. Perhaps we can look forward to added papers on this subject, where these many functions and Monte Carlo techniques will be applied to actual data derived from the sources under study, such as hospitalization claims under group insurance contracts or results of actual reinsurance treaties. Perhaps each problem requires a unique road of its own.

(AUTHOR'S REVIEW OF DISCUSSION)

NEWTON L. BOWERS, JR.:

I wish to thank Dr. Kahn, Dr. Berg, and Mr. Isen for their discussions of my paper. Their discussions bring up some of the problems of risk theory and a new solution to one of these problems.

Dr. Kahn in his discussion draws attention to a new paper by Kauppi and Ojantakanen, presented at the 1966 ASTIN Colloquium. The paper describes a method of using the standard normal distribution to approximate closely the well-known Esscher approximation. The technique involves a simple change of variable which takes account of the third moment. I highly recommend this paper to the interested reader.

Dr. Berg brings up a point about a typographical error which I would call a difference in notational preference. I use capital letters, Y, to indi-

[14] *Ibid.*, p. 442.

[15] Hugh H. Wolfenden, *The Fundamental Principles of Mathematical Statistics* (The Actuarial Society of America, 1942).

[16] *Ibid.*, p. 312.

cate the random variable and small letters, y, to indicate numerical values that the random variable assumes. While this notation is not universal, it has been used in several recently published textbooks.

Dr. Berg also shows a quadrature interpretation of the stop-loss net premium where the distribution of total claim amount is assumed to be a gamma distribution. A general interpretation of this type is possible. Looking at the integral for the stop-loss net premium, we see

$$\Pi(x) = \int_x^\infty (y-x)f(y)\,dy = \int_x^\infty [(y-\mu)-(x-\mu)]f(y)\,dy$$

$$= \int_x^\infty (y-\mu)f(y)\,dy - (x-\mu)[1-F(x)].$$

If the gamma density $f(y) = 1/\Gamma(\alpha)e^{-y}y^{\alpha-1}$ is substituted, the integral term may be evaluated as $x\,f(x)$, as Dr. Berg has pointed out. If $f(y)$ is instead the standard normal density, the integral term may be shown to equal simply $f(x)$.

Mr. Isen's comments raise some interesting questions, and I will answer them by giving my view of risk theory and by restating what I consider to be the place of my paper.

As I see it, there are two problems in risk theory. The first problem is to construct the probability model. There are two main classes of models. The first type, the collective model, divides the problem of calculating the distribution of total claim amounts into two stages. First a distribution of the number of claims is calculated. Certain assumptions, outlined in Dr. Kahn's paper, lead one to a Poisson distribution for the number of claims. Mr. Isen discusses these assumptions but seems to object to the stationarity assumption for the collective model on the basis of an analysis pertaining to an individual life. In any case, I agree with him that the Poisson assumptions seem inappropriate for small groups. Other actuaries have noted that in certain branches of the insurance industry the negative binomial distribution seems to fit the number of claims much more closely. The negative binomial distribution can, in fact, be obtained on theoretical grounds under the assumption that the Poisson parameter itself is a random variable with a gamma distribution.

The second part of the construction of the collective model is to obtain a probability distribution on the amount of a single claim. This is a problem which seems to me to be quite difficult and one which must be faced in casualty, health, and disability coverages. In these fields of insurance a statistical analysis of the distribution of individual claim sizes would be necessary. This analysis is then used to construct the so-called secondary distribution, the form of which is often assumed to be log-normal or

Pareto or gamma. It is possible in some of the so-called distribution-free methods to assume nothing about the distribution of individual claims but simply to use the data on claims to evaluate moments. If the data are sparse, the sample may give misleading results as to the distribution of large claim amounts.

The second model in general use is the individual model. The total claim amount is viewed as the sum of the claims on the various individual lives in the portfolio. The distribution of total claim amounts is the distribution of a sum of a fixed number of independent random variables. For life insurance where the amount of claim is, in general, fixed given the time of claim, this individual approach appears to be the preferred one.

Once the model is constructed, the second problem comes into play. How does one use the model to evaluate probabilities and expectations? It is virtually impossible to evaluate these probabilities directly because of difficulties in evaluating convolutions of functions.

There are several solutions to this second problem. One of the most popular methods at present is to use Monte Carlo techniques of simulation. If a large number of simulations are used, details of the distribution can be evaluated. Modern computers make such a scheme quite feasible.

Another method is to use the probability model to evaluate moments of the distribution of total claim amounts. A standard curve is then fitted to these moments and used to evaluate probabilities. Since the normal distribution and the gamma distribution are two parameter families of distributions, the values of these parameters may be chosen so that the mean and variance of the approximating probability curve are equal to those of the underlying model. If the third moment is known, the Pearson Type III curve may be used. Bohman and Esscher have shown that this last distribution gives much better results than the normal distribution for details of the tails of the distribution. The Gram-Charlier series and the Edgeworth series both allow one to use even more additional information in the form of knowledge of additional moments to improve the approximating density. The series suggested in my paper is like the Gram-Charlier series in that it makes use of additional moments. However, the underlying distribution used in generating the series is the gamma distribution rather than the normal distribution, and the results of the Bohman-Esscher paper suggested to me that this gamma series might give better results.

Another solution to the problem of evaluating various probability statements is by the use of the Esscher approximation. This involves evaluating moments of a special distribution function containing a multiplicative factor e^{hx}. It is a little unhandy to deal with, but it does give good results. Furthermore, the paper mentioned by Dr. Kahn shows a very good and very easy approximation to the Esscher method, which does not involve the evaluation of moments of the special distribution function.

Again I would like to thank those who took the time to prepare discussions of my paper and to allow me to indicate where I feel it fits into risk theory.