TRANSACTIONS OF SOCIETY OF ACTUARIES
1965 VOL. 17 PT. 1 NO. 49

EXCESS RATIO DISTRIBUTIONS IN RISK THEORY

DWIGHT K. BARTLETT, III

INTRODUCTION

A BASIC problem in risk theory is the calculation of expected losses in excess of some stated limit. This problem has its practical significance in experience rating of group policies and the calculation of premiums for stop loss or nonproportional reinsurance and similar problems. The term "excess ratio" has been adopted in this paper because this problem has been studied at some length by casualty actuaries and this is the terminology commonly used by them. The concern of casualty actuaries for this subject is primarily in the experience rating of Workmen's Compensation risks. In studying this problem the casualty actuaries have the benefit of a great many data which have been collected and can be used as the basis for determining the frequency distribution of the total losses for a particular risk in a given period of time. These data were used to construct a graduated frequency distribution called Table "M" some years ago and this table has been revised from time to time as more recent data have become available.

Unfortunately, life actuaries concerned with this problem do not have available the equivalent kind of data relating to the kinds of coverages they are used to dealing with. Several papers in the *Transactions* have suggested the use of the Monte Carlo simulation as a means for generating this kind of information, particularly with relation to group life coverage. Also, the paper "Experience Rating" by Paul H. Jackson [5],[1] and the discussions of that paper give some of the mathematical development of excess ratios, as well as some actual tables of values of excess ratios for group life coverage, which are calculated on the basis of a number of assumptions. In a recent paper "An Introduction to Collective Risk Theory and Its Application to Stop Loss Reinsurance" by Paul M. Kahn [6], the author extends the mathematical development of the determination of the frequency distribution of the aggregate losses for a group. Another fine reference in this general subject is the Society's *Study Note* on the subject of risk theory by John C. Wooddy [12], which summarizes the major developments in this field by life actuaries up to the present time. A study of these references will reveal the problems an actuary faces in the determination of excess ratios in practical situations in view of the

[1] Numbers in brackets refer to the bibliography at the end of this paper.

absence of worthwhile historical data. The purpose of this paper is to suggest methods for calculating excess ratios in the practical situation without the laborious calculations which the techniques of Dr. Kahn's paper would seem to require, and as an alternative to the Monte Carlo simulation technique which would also seem to be impractical to use on a case by case basis in experience rating. It further attempts to give some insight into the errors implicit in assuming $f(x)$ is one of the standard distributions, as is sometimes done.

I. COMPOUND POISSON PROCESS

The excess ratio function in mathematical terms is

$$\phi(n) = \int_n^\infty (x-n) f(x)\, dx \qquad \text{(for continuous distributions)}$$

or

$$= \sum_{x=n}^{\infty} (x-n) p(x) \qquad \text{(for discrete distributions),}$$

where $f(x)$ is a frequency distribution of aggregate losses[2] and $(x - n)$ is the aggregate losses in excess of a stated limit, n. In this section we will examine the statistical nature of the distribution $f(x)$. As pointed out in several of the previously mentioned references, $f(x)$ is a compound distribution resulting from the compounding of the probability distribution of the aggregate number of claims with the probability distribution of the size of a particular claim given that a claim has occurred, which we will also call hereafter the secondary distribution. In these references it has been assumed that the probability distribution of the number of claims is the result of a stochastic process which meets the conditions of a Poisson process and which can therefore be represented by the Poisson distribution.

The conditions of the Poisson process are twofold. The first is that in a small period of time, Δt, the probability of an event occurring is $\lambda \Delta t$ and that the probability of more than one event occurring is of such a small magnitude that it may be ignored in the mathematical analysis. The second condition for a Poisson process is that the parameter λ, which may be thought of as the rate at which events occur per unit of time, remain constant in the period of time in which the Poisson process is being studied. In the kinds of practical problems to which collective risk theory is applied these conditions seem to be reasonably well met. By the choice of a sufficiently small Δt the first condition is met, since it is highly unlikely

[2] Hereafter $f(x)$ is meant to symbolize the distribution of aggregate losses, regardless of whether the distribution is discrete or continuous.

that two claims would occur in exactly the same instance regardless of the number of claims expected in the policy year or calendar year under study. The second condition, the constancy of λ, is apt to prove somewhat more troublesome since it is a well-accepted fact that the rate of claims is higher on both life and accident and health coverages in the winter months than in the summer months. In addition, other factors affecting the rate of claims may change such as the age distribution, sex distribution, number of insured lives, etc. However, in the usual practical problem these variations do not appear large enough to invalidate the assumption of the second condition. There are a number of standard tests such as the chi-square test for the uniformity of events, and the Lexis Ratio test for the goodness of fit of the intervals between events to the negative exponential distribution, which may be used in testing the validity of this assumption. In only the most unusual circumstances do these tests indicate that the assumption of the constancy of λ is significantly violated. Therefore, in the rest of this paper it will be assumed that the probability distribution of the number of claims can be represented by the Poisson distribution with an appropriately chosen λ. However, it would be well to keep in mind the stated conditions for a Poisson process.

In general, the probability distribution of the size of claims cannot be assumed to follow any known distribution and this is where the main difficulty lies in finding $f(x)$. However, the moments of this distribution can usually be found in practical situations. This is all that need be known in order to find the moments of $f(x)$, as is demonstrated in the Appendix using the probability generating function technique. This is in general true for compound Poisson distributions. This allows us to proceed, as is shown in the following section, without requiring us to make any simplifying assumptions as to the mathematical nature of the probability distribution of the size of the claims.

II. JUSTIFICATION FOR THE APPROXIMATION OF $f(x)$ BY A GAMMA VARIATE AND EXAMINATION OF THE ERRORS THEREIN

We have shown in the Appendix that the moments of the compound Poisson distribution can be calculated from the moments of the secondary distribution (hereafter designated by μ_n^A) which in practical problems can usually be calculated. Although the complete set of moments of a distribution uniquely characterizes that distribution, it is unfortunately true that the distribution itself cannot be explicitly determined in general from the moments. The technique adopted here involves choosing a distribution function which seems intuitively to have the general shape of the compound Poisson distribution and then comparing the moments of the two distributions.

Of all the probability distributions in common use, the one which would intuitively appear to represent reasonably the frequency distribution of aggregate losses on small groups of insured lives is the gamma distribution. This function is

$$f(x) = \frac{1}{\alpha!\,\beta^{\alpha+1}}\, x^{\alpha} e^{-x/\beta} \qquad \begin{matrix} (x \geq 0) \\ (\alpha > -1) \\ (\beta > 0). \end{matrix}$$

Figure 1 shows the general shape of this distribution for $\beta = 1$ and various values of α. This function has the virtue that as α increases it tends

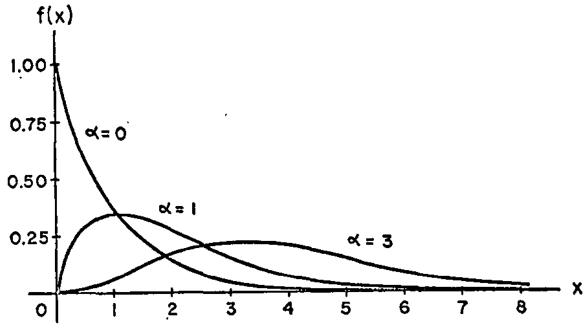

FIG. 1

to approach the normal distribution in the limit, which would seem reasonable as the assumed frequency distribution of aggregate claims on larger size groups. The moments of this function are as follows:

$$\mu_1 = \beta(\alpha + 1)\,,$$

$$\mu_2 = \beta^2(\alpha + 1)\,,$$

$$\mu_3 = 2\beta^3(\alpha + 1)\,,$$

$$\mu_4 = 3\beta^4(\alpha + 1)(\alpha + 3),\ \text{etc}.$$

It is the position of this paper that the frequency distribution of the aggregate amount of claims, $f(x)$, can be reasonably approximated by the gamma distribution in most practical cases if the parameters α and β are determined by setting the expressions for the first two moments of the gamma distribution equal to the first two moments of the compound Poisson distribution as given in the Appendix.

$$\beta(\alpha + 1) = \lambda\pi'(1) = \lambda\mu_1^A,$$

$$\beta^2(\alpha + 1) = \lambda(\pi''(1) + \pi'(1)) = \lambda(\mu_2^A + (\mu_1^A)^2),$$

where μ_1^A and μ_2^A are the mean and variance, respectively, of the secondary distribution.

To illustrate the suggestion and to give some insight into the kind of error that it involves, let us assume that the secondary distribution is a geometric distribution, $P(x = k) = pq^k$ for $k = 0, 1, 2, \ldots$, with p.g.f. $\pi(s) = P/(1 - qs)$. By differentiating with respect to s and letting $s = 1$, we find

$$\pi'(1) = q/p,$$

$$\pi''(1) = 2q^2/p^2,$$

$$\pi'''(1) = 6q^3/p^3,$$

and

$$\pi''''(1) = 24q^4/p^4.$$

Let

$$\beta(\alpha + 1) = \lambda\pi'(1) = \lambda q/p$$

and

$$\beta^2(\alpha+1) = \lambda(\pi''(1) + \pi'(1)) = \lambda\left(\frac{2q^2}{p^2} + \frac{q}{p}\right).$$

Solving for α and β, we find

$$\beta = (\pi''(1) + \pi'(1))/\pi'(1) = \frac{2q}{p} + 1$$

and

$$\alpha + 1 = \lambda(\pi'(1))^2/(\pi''(1) + \pi'(1)) = \frac{\lambda q}{2 - p}.$$

If we now examine the third moments, we find that for the gamma function

$$\mu_3 = 2\beta^3(\alpha + 1) = \frac{2(\pi''(1) + \pi'(1))^2\lambda}{\pi'(1)}$$

$$= \lambda\left(\frac{8q^3}{p^3} + \frac{8q^2}{p^2} + \frac{2q}{p}\right).$$

However, using the results of the Appendix, we can show

$$\mu_3 = \lambda(\pi'''(1) + 3\pi''(1) + \pi'(1))$$

$$= \lambda\left(\frac{6q^3}{p^3} + \frac{6q^2}{p^2} + \frac{q}{p}\right)$$

for the particular compound Poisson process.

If we take the ratio of the third moments, we have a ratio of the skewnesses of the two distributions, since the generally used measure of skewness is $\mu_3/\mu_2^{3/2}$. The ratio of the skewnesses of the gamma distribution and the particular compound Poisson distribution necessarily lies between $\frac{4}{3}$ and 2, depending on the value of q/p. Since the skewness of the gamma distribution is the larger of the two, it appears that it has a larger area under the right-hand tail and would therefore tend to overstate the excess ratio. However, it seems likely that the order of error is less than the ratio of the third moments, since the third moments involve x^3, while the excess ratio involves x.

Proceeding similarly with the fourth moments, we find for the gamma distribution that

$$\begin{aligned}\mu_4 &= 3\beta^4(\alpha+1)(\alpha+3)\\ &= 3\lambda^2(\pi''(1)+\pi'(1))^2+\frac{6\lambda(\pi''(1)+\pi'(1))^3}{[\pi'(1)]^2}\\ &= 12\lambda^2\frac{q^2}{p^2}\left(\frac{q}{p}+1\right)^2+48\lambda\frac{q^4}{p^4}+72\lambda\frac{q^3}{p^3}+36\lambda\frac{q^2}{p^2}+\frac{6\lambda q}{p}.\end{aligned}$$

Using the results of the Appendix, we find for the particular compound Poisson distribution

$$\begin{aligned}\mu_4 &= 3\lambda^2(\pi''(1)+\pi'(1))^2+\lambda(\pi''''(1)+6\pi'''(1)+7\pi''(1)+\pi'(1))\\ &= 12\lambda^2\frac{q^2}{p^2}\left(\frac{q}{p}+1\right)^2+24\lambda\frac{q^4}{p^4}+36\lambda\frac{q^3}{p^3}+14\lambda\frac{q^2}{p^2}+\lambda\frac{q}{p}.\end{aligned}$$

The ratio of the fourth moments is the ratio of the kurtosis of the distributions, since the generally used measure of kurtosis is μ_4/μ_2^2. It can be seen that, if λ is large so that the term in the kurtosis involving λ^2 predominates, the ratio will be close to 1. If λ is small and q/p is large, the ratio will approach 2. If λ and q/p are both small, the ratio will tend to 6. In most practical situations q/p would tend to be large if we think of p as the probability of the termination of a disability claim in a particular day or something similar. Therefore, the ratio will normally be in the range of 1 and 2. The exact significance of this is hard to assess, but it would seem to mean that there was relatively more area of the curve under the distance tails in the gamma function than in the particular compound Poisson function, which would again mean that the gamma function gives an overstatement of the excess ratio.

The assumption that the frequency distribution of a particular claim is a geometric distribution would seem a worse possible situation, since it admits of a theoretically infinite size maximum claim. It is instructive to examine the situation when the secondary distribution has a maximum claim more realistically related to the average size claim. To do this, let us assume that the secondary distribution is a uniform distribution with $\mu_1^A = a$ and with a range from 0 to $2a$. The probability distribution is therefore

$$P(x = k) = \frac{1}{2a+1} \quad \text{for} \quad k = 0, 1, 2, \ldots, 2a.$$

The probability generating function is

$$\pi(s) = \frac{1}{(2a+1)}(1 + s + s^2 + \ldots, + s^{2a})$$

$$= \frac{1}{2a+1}\left(\frac{1 - s^{2a+1}}{1 - s}\right).$$

Therefore

$$\pi'(s) = \sum_{n=0}^{2a} n\, s^{n-1} \Big/ (2a+1)$$

$$\pi'(1) = a$$

$$\pi''(s) = \sum_{n=0}^{2a} n(n-1)\, s^{n-2} \Big/ (2a+1)$$

$$\pi''(1) = \frac{(2a-1)\,2a}{3}$$

$$\pi'''(s) = \sum_{n=0}^{2a} n(n-1)(n-2)\, s^{n-3} \Big/ (2a+1)$$

$$\pi'''(1) = \frac{(2a-2)(2a-1)\,2a}{4}.$$

Using the results of the Appendix, we determine therefore that

$$\mu_1 = \lambda a$$

$$\mu_2 = \lambda\left(\frac{4a^2 + a}{3}\right).$$

Proceeding as with the geometric distribution, let

$$\mu_1 = \beta(\alpha + 1) = \lambda a$$

$$\mu_2 = \beta^2(\alpha + 1) = \lambda\left(\frac{4a^2 + a}{3}\right)$$

$$\therefore \beta = \frac{4a+1}{3} \quad \text{and} \quad \alpha + 1 = \frac{3\lambda a}{4a+1}.$$

The third moment for the gamma distribution will therefore be

$$\mu_3 = 2\beta^3(\alpha + 1) = \lambda\left(\frac{32a^3 + 16a^2 + 2a}{9}\right).$$

The third moment for the particular compound Poisson distribution is

$$\begin{aligned}\mu_3 &= \lambda(\pi'''(1) + 3\pi''(1) + \pi'(1)) \\ &= \lambda(2a^3 + a^2)\,.\end{aligned}$$

The ratio of the third moment, and therefore the ratio of the skewnesses, will tend to $\frac{16}{9}$ as a becomes large. Note that the third moment for the gamma distribution is larger than for the particular compound Poisson distribution, suggesting again that the use of the gamma distribution would tend to overstate the excess ratio.

III. EXAMINATION OF THE ERRORS IN APPROXIMATION OF $f(x)$ BY A SIMPLE POISSON DISTRIBUTION

If it is assumed that $f(x)$ can be represented by a simple Poisson distribution (i.e., the secondary distribution is a "spike" distribution with $\mu_2^A = 0$), it can be seen from the preceding results that the errors can be very substantial. For example, if the secondary distribution is really a geometric, we have seen that μ_2 for the compound Poisson distribution is

$$\lambda\left(\frac{2q^2}{p^2} + \frac{q}{p}\right).$$

However, for the simple Poisson $\mu_1 = \mu_2$, and consequently if the mean number of claims is λ, the variance in the amount of claims is $\lambda(q/p)^2$ under this assumption. If q/p is large, the error in the variance will approach 100 per cent.

Similarly we have seen that μ_2 for the uniform distribution is

$$\lambda\left(\frac{4a^2 + a}{3}\right).$$

Using the simple Poisson would give a variance in amount of claims of λa^2. Therefore the error in the variance, in this case, would approach $33\frac{1}{3}$ per cent as a became large.

IV. THE APPROXIMATION OF $f(x)$ BY THE NORMAL DISTRIBUTION

Examination of the skewness of the compound Poisson distribution gives insight into when $f(x)$ can reasonably be approximated by a normal distribution. The skewness of the general compound Poisson distribution, using the results of the Appendix, is seen to be

$$\frac{\mu_3}{(\mu_2)^{3/2}} = \frac{\lambda(\pi'''(1) + 3\pi''(1) + \pi'(1))}{\lambda^{3/2}(\pi'(1) + \pi''(1))^{3/2}},$$

but the moments of the secondary distribution are as follows:

$$\mu_1^A = \pi'(1)$$

$$\mu_2^A = \pi''(1) + \pi'(1) - [\pi'(1)]^2$$

$$\mu_3^A = \pi'''(1) + [1 - \pi'(1)][3\pi''(1) + \pi'(1) - 2(\pi'(1))^2] .$$

By successive substitutions we find

$$\pi'(1) = \mu_1^A$$

$$\pi''(1) = \mu_2^A - \mu_1^A + (\mu_1^A)^2$$

$$\pi'''(1) = \mu_3^A - (1 - \mu_1^A)[3\mu_2^A - 2\mu_1^A + (\mu_1^A)^2] .$$

Substitution in (1) gives

$$\frac{\mu_3}{(\mu_2)^{3/2}} = \frac{\mu_3^A + 3\mu_2^A\mu_1^A + (\mu_1^A)^3}{\lambda^{1/2}[\mu_2^A + (\mu_1^A)^2]^{3/2}} .$$

If we assume that the secondary distribution is "standardized" so that

$$\mu_1^A = 0 \quad \text{and} \quad \mu_2^A = 1 ,$$

then

$$\frac{\mu_3}{(\mu_2)^{3/2}} = \frac{\mu_3^A}{\lambda^{1/2}} .$$

In a standardized asymmetrical distribution μ_3^A might typically be of the order of 1 to 2. The skewness of the normal distribution is 0, since it is symmetrical. By the time λ has reached 100, the skewness of the compound Poisson distribution would be of the order of .1 to .2, which would seem small enough to make the assumption that $f(x)$ was a normal distribution reasonable. In any case resort to the normal distribution will have to be made at some point since tables of the incomplete gamma function, which are required, as will be seen later, if $f(x)$ is assumed to be a gamma distribution, simply do not exist beyond a certain size of the parameters.

When the normal approximation is used, the suggested procedure is the same as for the gamma, that is, setting the first two moments of the particular compound Poisson distribution equal to the first two moments of the normal distribution. This will give the parameters for the normal distribution, since the parameters of the normal distribution are the first two moments

$$\mu_1^N = \lambda\mu_1^A$$

and

$$\mu_2^N = \lambda(\mu_2^A + (\mu_1^A)^2) .$$

V. THE SOLUTION OF THE EXCESS RATIO FUNCTION FOR STANDARD DISTRIBUTIONS

One of the fortuitous characteristics of the excess ratio function

$$\phi(n) = \int_n^\infty (x-n) f(x)\, dx$$

is that it has a solution which permits the use of generally available tables, if $f(x)$ is one of the common frequency or probability distributions. This section will demonstrate the solution for a number of the distributions commonly used in computing excess ratios.

We will solve the excess ratio function for the Poisson distribution first.

$$\phi(n) = \sum_{x=n}^{\infty} (x-n) \frac{e^{-\lambda}(\lambda)^x}{x!}$$

$$= \sum_{x=n}^{\infty} \frac{x e^{-\lambda}(\lambda)^x}{x!} - n \sum_{x=n}^{\infty} \frac{e^{-\lambda}(\lambda)^x}{x!}$$

$$= \lambda \sum_{x=n}^{\infty} \frac{e^{-\lambda}(\lambda)^{x-1}}{(x-1)!} - n \sum_{x=n}^{\infty} \frac{e^{-\lambda}(\lambda)^x}{x!} .$$

Let

$$P(n, \lambda) = \sum_{x=0}^{n} \frac{e^{-\lambda}\lambda^x}{x!} ,$$

the cumulative distribution. Then

$$\phi(n) = \lambda(1 - P(n-2, \lambda)) - n(1 - P(n-1, \lambda)) .$$

The solution of $\phi(n)$ for the gamma distribution is as follows:

$$\phi(n) = \int_n^\infty (x-n) \frac{1}{\alpha!\, \beta^{\alpha+1}} x^\alpha e^{-x/\beta} dx$$

$$= \int_n^\infty \frac{1}{\alpha!\, \beta^{\alpha+1}} x^{\alpha+1} e^{-x/\beta} dx - n \int_n^\infty \frac{1}{\alpha!\, \beta^{\alpha+1}} x^\alpha e^{-x/\beta} dx$$

$$= \beta(\alpha+1) \int_n^\infty \frac{1}{(\alpha+1)!\, \beta^{\alpha+2}} x^{\alpha+1} e^{-x/\beta} dx$$

$$- n \int_n^\infty \frac{1}{\alpha!\, \beta^{\alpha+1}} x^\alpha e^{-x/\beta} dx .$$

Let

$$F(n; \alpha, \beta) = \int_0^n \frac{1}{\alpha!\, \beta^{\alpha+1}} x^\alpha e^{-x/\beta} dx ,$$

the cumulative distribution

$$\phi(n) = \beta(\alpha+1)[1 - F(n; \alpha+1, \beta)] - n[1 - F(n; \alpha; \beta)] .$$

The solution of $\phi(n)$ for the normal distribution is as follows:

$$\phi(n)=\frac{1}{(2\pi\sigma^2)^{1/2}}\int_n^\infty (x-n)\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx$$

$$=\frac{1}{(2\pi\sigma^2)^{1/2}}\int_{x=n}^\infty (x-\mu)\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)d(x-\mu)$$

$$-\frac{(n-\mu)}{(2\pi\sigma^2)^{1/2}}\int_n^\infty \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx$$

$$=\frac{\sigma}{(2\pi)^{1/2}}\int_{x=n}^\infty \exp-\frac{(x-\mu)^2}{2\sigma^2}\,d\,\frac{(x-\mu)^2}{2\sigma^2}$$

$$-\frac{(n-\mu)}{(2\pi\sigma^2)^{1/2}}\int_n^\infty \exp-\frac{(x-\mu)^2}{2\sigma^2}\,dx$$

$$=\frac{\sigma}{(2\pi)^{1/2}}\left[\exp-\frac{(x-\mu)^2}{2\sigma^2}\right]_\infty^n-\frac{(n-\mu)}{(2\pi\sigma^2)^{1/2}}\int_n^\infty \exp-\frac{(x-\mu)^2}{2\sigma^2}\,dx$$

$$=\frac{\sigma}{(2\pi)^{1/2}}\left(\exp-\frac{(n-\mu)^2}{2\sigma^2}\right)-(n-\mu)(1-F(n;\mu,\sigma))$$

where

$$F(n;\mu,\sigma)=\frac{1}{(2\pi\sigma^2)^{1/2}}\int_{-\infty}^n \exp-\frac{(x-\mu)^2}{2\sigma^2}\,dx\,.$$

The binomial and other distributions can similarly be solved.

In Table 1 are set forth actual values of $\phi(n)$ for the gamma distribution

TABLE 1

VALUES OF $\phi(n)\cdot\mu_1^A/[\mu_2^A+(\mu_1^A)^2]$

(Y/X)	GAMMA (X)									NORMAL (X)
	0.1	0.3	0.5	1.0	2.5	5.0	10.0	25.0	50.0	50.0
100%..	0.075	0.172	0.242	0.368	0.610	0.876	1.250	1.988	2.814	2.821
110...	0.074	0.165	0.226	0.333	0.513	0.677	0.852	1.042	1.057	0.999
120...	0.072	0.157	0.212	0.301	0.431	0.518	0.564	0.495	0.313	0.251
130...	0.071	0.150	0.199	0.273	0.360	0.392	0.362	0.215	0.075	0.042
140...	0.069	0.143	0.187	0.247	0.300	0.292	0.225	0.086	0.015	0.005
150...	0.068	0.137	0.175	0.223	0.249	0.216	0.136	0.032	0.002	0.000+
160...	0.066	0.131	0.165	0.202	0.207	0.159	0.081	0.011	0.000+	0.000+
170...	0.065	0.125	0.155	0.183	0.171	0.117	0.047	0.003	0.000+	0.000+
180...	0.064	0.120	0.146	0.166	0.158	0.084	0.027	0.001	0.000+	0.000+
190...	0.062	0.115	0.137	0.150	0.116	0.060	0.015	0.000+	0.000+	0.000+
200...	0.061	0.110	0.129	0.136	0.095	0.042	0.008	0.000+	0.000+	0.000+

or, more accurately speaking,

$$\frac{\phi(n)\cdot\mu_1^A}{\mu_2^A+(\mu_1^A)^2}.$$

Setting forth the table in this form permits its reduction to two dimensions. Recall that for the gamma function

$$\phi(n)=\beta(a+1)\int_n^\infty \frac{x^{a+1}e^{-x/\beta}}{(a+1)!\,\beta^{a+2}}\,dx-n\int_n^\infty \frac{x^a e^{-x/\beta}}{a!\,\beta^{a+1}}\,dx$$

$$=\beta(a+1)\int_{n/\beta}^\infty\left(\frac{x}{\beta}\right)^{a+1}\frac{e^{-x/\beta}}{(a+1)!}\frac{dx}{\beta}-n\int_{n/\beta}^\infty\left(\frac{x}{\beta}\right)^a\frac{e^{-x/\beta}}{a!}\,d\,\frac{x}{\beta}.$$

Let

$$I\left(\frac{n}{\beta},\ a+1\right)=\int_0^{n/\beta}\left(\frac{x}{\beta}\right)^{a+1}\frac{e^{-x/\beta}}{(a+1)!}\,d\,\frac{x}{\beta}.$$

Therefore

$$\phi(n)=\beta(a+1)\left[1-I\left(\frac{n}{\beta},\ a+1\right)\right]-n\left[1-I\left(\frac{n}{\beta},\ a\right)\right].$$

Recall from Section II that

$$\beta=\frac{\pi''(1)+\pi'(1)}{\pi'(1)}=\frac{\mu_2^A+(\mu_1^A)^2}{\mu_1^A}$$

and

$$a+1=\frac{\lambda(\pi'(1))^2}{\pi''(1)+\pi'(1)}=\frac{\lambda(\mu_1^A)^2}{\mu_2^A+(\mu_1^A)^2}.$$

Then

$$\phi(n)=\lambda\mu_1^A\left[1-I\left(\frac{n\mu_1^A}{\mu_2^A+(\mu_1^A)^2},\quad \frac{\lambda(\mu_1^A)^2}{\mu_2^A+(\mu_1^A)^2}\right)\right]$$

$$-n\left[1-I\left(\frac{n\mu_1^A}{\mu_2^A+(\mu_1^A)^2},\quad \frac{\lambda(\mu_1^A)^2}{\mu_2^A+(\mu_1^A)^2}-1\right)\right].$$

$$\frac{\phi(n)\cdot\mu_1^A}{\mu_2^A+(\mu_1^A)^2}=\frac{\lambda(\mu_1^A)^2}{\mu_2^A+(\mu_1^A)^2}\left[1-I\left(\frac{n\mu_1^A}{\mu_2^A+(\mu_1^A)^2},\quad \frac{\lambda(\mu_1^A)^2}{\mu_2^A+(\mu_1^A)^2}\right)\right]$$

$$-\frac{n\mu_1^A}{\mu_2^A+(\mu_1^A)^2}\left[1-I\left(\frac{n\mu_1^A}{\mu_2^A+(\mu_1^A)^2},\quad \frac{\lambda(\mu_1^A)^2}{\mu_2^A+(\mu_1^A)^2}-1\right)\right].$$

Let

$$X = \frac{\lambda(\mu_1^A)^2}{\mu_2^A + (\mu_1^A)^2}$$

and

$$Y = \frac{n\mu_1^A}{\mu_2^A + (\mu_1^A)^2} = \frac{n}{\lambda\mu_1^A}(X).$$

$$\frac{\phi(n)\mu_1^A}{\mu_2^A + (\mu_1^A)^2} = X[1 - I(Y, X)] - Y[1 - I(Y, X - 1)].$$

The table is set forth with the entries X and Y as thus determined and Y, the "stop-loss" level, expressed as a percentage of X. The percentage can be also thought of as the ratio of the stop-loss level to the pure risk premium. Since

$$\frac{\mu_1^A}{\mu_2^A + (\mu_1^A)^2} = \frac{\mu_1}{\mu_2},$$

the ratio of the first and second moments of $f(x)$, the table can also be used if the moments of $f(x)$ are known without reference to the moments of the secondary distribution. In this case

$$\phi(n)\left(\frac{\mu_1}{\mu_2}\right) = \frac{\theta\mu_1}{\mu_2}\left[1 - I\left(\frac{n\mu_1}{\mu_2}, \frac{\theta\mu_1}{\mu_2}\right)\right] - \frac{n\mu_1}{\mu_2}\left[1 - I\left(\frac{n\mu_1}{\mu_2}, \frac{\theta\mu_1}{\mu_2} - 1\right)\right]$$

where θ is the pure risk premium.

To show the small difference in $\phi(n)$ if $f(x)$ is a normal or a gamma distribution with the same first two moments, $\phi(n)$ is tabulated on both assumptions for $X = 50$. Note that the gamma distribution gives larger values of $\phi(n)$ due to its positive skewness.

VI. RISK CHARGES

In the experience rating of group insurance, it is usual to "forgive" losses in any year in excess of some limit and not to charge these excess claims against the dividend rating of subsequent years. The excess ratio function tells us on the average how much in claims has to be forgiven. However, how this "forgiveness" is to be funded is another problem. Mr. Jackson, in his paper previously referred to [5], suggests two possible methods. The first is to charge the expected excess as a level charge against each policy without reference to its particular claim experience. The other, which he calls the "J" method is to charge each case with a fixed percentage of its claim profit, that is, the amount available to pay claims less the actual cost of claims.

In the notation of this paper the claim profit function would be:

$$\sigma(m)=\int_{-\infty}^{m}(m-x)f(x)dx .$$

The function, $\sigma(m)$, can be solved in a very similar manner to the solution of $\phi(n)$. Table 2 tabulates $\sigma(m)$ in the same general way as Table 1 tabulates $\phi(n)$. The "stop-loss level" in this case is the total premium available to pay claims as a percentage of the pure risk premium. The two

TABLE 2

VALUES OF $\sigma(m)\cdot\mu_1^A/[\mu_2^A+(\mu_1^A)^2]$

(Y/X)	GAMMA (X)									NORMAL (X)
	0.1	0.3	0.5	1.0	2.5	5.0	10.0	25.0	50.0	50.0
100%..	0.075	0.172	0.242	0.368	0.610	0.876	1.250	1.988	2.814	2.821
110....	0.084	0.195	0.276	0.433	0.763	1.177	1.852	3.542	6.057	5.999
120....	0.092	0.217	0.312	0.501	0.931	1.518	2.564	5.495	10.313	10.250
130....	0.101	0.240	0.349	0.573	1.110	1.892	3.362	7.715	15.075	15.040
140....	0.109	0.263	0.387	0.647	1.300	2.292	4.225	10.086	20.015	20.000
150....	0.118	0.287	0.425	0.723	1.499	2.716	5.136	12.532	25.002	25.003
160....	0.126	0.311	0.465	0.802	1.707	3.159	6.081	15.011	30.000	30.000
170....	0.135	0.335	0.505	0.883	1.921	3.617	7.047	17.503	35.000	35.000
180....	0.144	0.360	0.546	0.966	2.158	4.084	8.027	20.001	40.000	40.000
190....	0.152	0.385	0.586	1.050	2.366	4.560	9.015	22.500	45.000	45.000
200....	0.161	0.410	0.629	1.136	2.595	5.043	10.008	25.000	50.000	50.000

levels, m and n, for a particular experience rating formula, may or may not be the same.

The risk charge, if method "J" is used, is the ratio of the appropriate Table 1 value divided by the appropriate Table 2 value times the claim profit on the particular group policy.

VII. PRACTICAL ILLUSTRATION

As has previously been shown, the techniques suggested by this paper require the calculation of the first two moments of the secondary distribution, the distribution of claims by size. This can be done for life insurance by the following formula:

$$\mu_1^A=\sum_{i=1}^{n}q_iA_i\Big/\sum_{i=1}^{n}q_i$$

$$\mu_2^A=\Big(\sum_{i=1}^{n}q_i(A_i)^2-(\mu_1^A)^2\sum_{i=1}^{n}q_i\Big)\Big/\sum_{i=1}^{n}q_i ,$$

where q_i and A_i are the rate of mortality and amount of life insurance at risk on the ith life.

This can be done for health coverages by sampling claims. If continuation tables for the coverage in question exist, however, they can be used to calculate the moments. To illustrate, let us use the continuation table for group weekly disability income insurance in Morton D. Miller's paper, "Group Weekly Indemnity Continuation Table," *TSA*, III, 48–49. Let us assume the coverage in question has a seven-day waiting period with a thirteen-week benefit and that all insured lives are covered for a benefit of \$1.00 a day for each day of disability. The probability of a claim lasting exactly t days beyond the waiting period is

$$p_{t \text{ days}} = \frac{\Delta l_{t+7 \text{ days}}}{l_{8 \text{ days}}}, \quad 1 \text{ day} \leq t \leq 28 \text{ days}$$

$$p_{t \text{ weeks}} = \frac{\Delta l_{t+1 \text{ week}}}{l_{8 \text{ days}}}, \quad 5 \text{ weeks} \leq t \leq 13 \text{ weeks}$$

$$p_{91 \text{ days}} = \frac{l_{15 \text{ weeks}}}{l_{8 \text{ days}}}.$$

Since the continuation table is given only for weeks after 35 days, beyond that point the probabilities are for a claim ending in a particular week. It is assumed in Table 3 that claims ending in a particular week end at the close of the third day of that week of disability. Therefore a claim ending in the fifth week is assumed to end after 31 days. For this plan

$$\mu_1^A = \sum_{t=1}^{91} t p_t = 31.353 \text{ days}$$

$$\mu_2^A = \sum_{t=1}^{91} t^2 p_t - (\mu_1^A)^2 = 878.69$$

$$\frac{\mu_2^A + (\mu_1^A)^2}{\mu_1^A} = 59.379.$$

The X entry needed for using Tables 1 and 2 is

$$X = \frac{\lambda (\mu_1^A)^2}{\mu_2^A + (\mu_1^A)^2}.$$

The value of λ per life insured can be determined by dividing the pure risk premium, $\lambda \mu_1^A$, for the coverage by the value of μ_1^A. Table XI of Mr.

Miller's paper indicates that the value is

$$\lambda \text{ per life per year} = 7(c_8 - c_7) = .1463 \text{ per life insured} .$$

Therefore $X = .07725$ times the average number of lives insured during the year. If we have 100 lives with a stop-loss level of 120 per cent of the pure risk premium, Table 1 shows that $\phi(120 \text{ per cent}) = (.543)(59.379) = 32.24$.

TABLE 3

PROBABILITY DISTRIBUTION OF GROUP WEEKLY DISABILITY INCOME CLAIMS BY LENGTH OF CLAIM

(Thirteen-Week Maximum Benefit, Seven-Day Waiting Period, No Maternity Benefit)

Length of Claim (in Days)	p_t	Length of Claim (in Days)	p_t
1	.03500	20	.01510
2	.03474	21	.01465
3	.03349	22	.01374
4	.03318	23	.01334
5	.03195	24	.01295
6	.03160	25	.01214
7	.03040	26	.01180
8	.03002	27	.01106
9	.02885	28	.01076
10	.02701	31	.06361
11	.02530	38	.04832
12	.02370	45	.03753
13	.02222	52	.02980
14	.02083	59	.02399
15	.01953	66	.01939
16	.01831	73	.01586
17	.01772	80	.01300
18	.01662	87	.01077
19	.01611	91	.12561
			1.00000

If the claim profit is measured against 110 per cent of the pure risk premium, Table 2 shows that $\sigma(110 \text{ per cent}) = (1.545)(59.379) = 91.74$. If the "J" type of risk charge formula is used the risk charge as a percentage of claims profit would be $(.543)/(1.545) = 35.1$.

CONCLUSION

The paper has examined excess ratio distributions in the rarefied atmosphere of purely mathematical analysis. The author recognizes the limitations of such an analysis which have been pointed out in other references on the subject. The analysis assumes that deviations from expected experience are purely a result of random fluctuations in experience. It does not take into consideration unusual influences affecting the expected claim experiences such as epidemics, wars and natural catastrophes.

The paper assumes that the expected experience is known. In the experience rating of group policies this assumption is not necessarily valid, particularly in the early policy years before the premium rates have had an opportunity to reflect the actual experience of the group.

Another problem the paper has not examined is the situation in which several different coverages are being experience rated as a single entity. It is common, for example, in group dividend formulae to experience rate for dividend purposes the life, weekly disability income, and medical care coverages in a combined package. If the profit under one particular group coverage is available to absorb the possible losses under another coverage, the risk charge should obviously be less than the sum of the risk charges calculated separately for the individual coverages.

In any case, the author would reiterate the advantage of having frequency distributions of aggregate losses determined from actual experience. However, the mathematical analysis of the paper can be useful in the absence of such distributions.

The author wishes to thank his staff for their help in the preparation of this paper and, in particular, the preparation of the tables.

APPENDIX

The mathematical expression of the Poisson distribution is

$$p(k', \lambda) = e^{-\lambda}\lambda^k/k!\,, \qquad \lambda > 0 \qquad \text{and} \qquad k = 0, 1, 2, \ldots .$$

The probability generating function of the Poisson distribution is

$$\begin{aligned} G(s) &= p(0; \lambda) + p(1; \lambda)s + p(2; \lambda)s^2 + \ldots + p(n; \lambda)s^n + \ldots \\ &= e^{-\lambda(1-s)}\,. \end{aligned}$$

The probability generating function of a compound distribution is found simply by substituting for s the probability generating function of the secondary distribution. Let us call the p.g.f. of the secondary distribution $\pi(s)$. Therefore, the p.g.f. of the compound Poisson distribution is

$$G(\pi(s)) = e^{-\lambda(1-\pi(s))}\,.$$

The moments of a distribution can be found in general from the derivatives of the p.g.f. valued for $s = 1$ as follows:

$$\mu_1 = G'(1)$$
$$\mu_2 = G''(1) + G'(1) - [G'(1)]^2$$
$$\mu_3 = G'''(1) + (1 - G'(1))[3G''(1) + G'(1) - 2(G'(1))^2]$$
$$\mu_4 = G''''(1) - (4\mu_1 - 6)\mu_3 - (6\mu_1^2 - 8\mu_1 + 11)\mu_2 - (\mu_1^4 - 6\mu_1^3 + 11\mu_1^2 - 6\mu_1) .$$

The derivatives of $G(\pi(s)) = e^{-\lambda(1-\pi(s))}$ are as follows:

$$G'(\pi(s)) = G(\pi(s))\lambda\pi'(s) ,$$
$$\therefore G'(\pi(1)) = G(\pi(1))\lambda\pi'(1) = \lambda\pi'(1) ,$$

since $G(\pi(1)) = 1$. Likewise

$$G''(\pi(s)) = G(\pi(s))(\lambda\pi'(s))^2 + G(\pi(s))\lambda\pi''(s)$$
$$\therefore \quad G''(\pi(1)) = (\lambda\pi'(1))^2 + \lambda\pi''(1)$$
$$G'''(\pi(s)) = G'(\pi(s))[(\lambda\pi'(s))^2 + \lambda\pi''(s)] + G(\pi(s))[2\lambda^2\pi'(s)\pi''(s) + \pi'''(s)]$$
$$\therefore \quad G'''(\pi(1)) = \lambda\pi'(1)[(\lambda\pi'(1))^2 + \lambda\pi''(1)] + 2\lambda^2\pi'(1)\pi''(1) + \lambda\pi'''(1)$$

and

$$G''''(\pi(s)) = G''(\pi(s))[(\lambda\pi'(s))^2 + \lambda\pi''(s)] + 2G'(\pi(s))[2\lambda^2\pi'(s)\pi''(s) + \lambda\pi'''(s)] + G(\pi(s)[2\lambda^2(\pi''(s))^2 + 2\lambda^2\pi'(s)\pi'''(s) + \lambda\pi''''(s)] .$$
$$\therefore G''''(\pi(1)) = [(\lambda\pi'(1))^2 + \lambda\pi''(1)]^2 + 2\lambda\pi'(1)[2\lambda^2\pi'(1)\pi''(1) + \lambda\pi'''(1)] + [2\lambda^2(\pi''(1))^2 + 2\lambda^2\pi'\pi'''(1) + \lambda\pi''''(1)] .$$

Substituting these relationships in the equations for the moments, we get

$$\mu_1 = \lambda\pi'(1)$$
$$\mu_2 = \lambda(\pi'(1) + \pi''(1))$$
$$\mu_3 = \lambda(\pi'''(1) + 3\pi''(1) + \pi'(1))$$
$$\mu_4 = 3\lambda^2[(\pi''(1) + \pi'(1))^2] + \lambda(\pi''''(1) + 6\pi'''(1) + 7\pi''(1) + \pi'(1)) .$$

Thus we see, as stated in Section I of the paper, that the moments of the compound Poisson distribution can be computed directly from the moments of the secondary distribution.

BIBLIOGRAPHY

1. BAILEY, NORMAN T. J. *The Elements of Stochastic Processes with Applications to the Natural Sciences.* New York: John Wiley & Sons, Inc., 1964.
2. COLLINS, RUSSELL M. "Actuarial Application of the Monte Carlo Technique," *TSA*, XIV (1962), 365.
3. FEAY, H. L. "Introduction to Nonproportional Reinsurance," *TSA*, XII (1960), 22.
4. FELLER, WILLIAM. *An Introduction to Probability Theory and Its Applications*, Vol. I. New York: John Wiley & Sons, Inc., 1950.
5. JACKSON, PAUL H. "Experience Rating," *TSA*, V (1953), 239.
6. KAHN, PAUL M. "An Introduction to Collective Risk Theory and Its Application to Stop Loss Reinsurance," *TSA*, XIV (1962), 400.
7. MOOD, ALEXANDER McF. *Introduction to the Theory of Statistics.* New York: McGraw-Hill Book Co., Inc., 1950.
8. PEARSON, K. *Tables of the Incomplete Gamma Functions.* London, 1922.
9. ROBERTS, LEWIS H. "Graduation of Excess Ratio Distributions by the Method of Moments," *PCAS*, XLIV (1957), 45.
10. SIMON, LEROY J. "The 1965 Table M." Paper presented to the Casualty Actuarial Society May 26, 1965, to be published later as part of *Proceedings of the Casualty Actuarial Society.*
11. VALERIUS, NELS M. "Risk Distributions Underlying Insurance Charges in the Retrospective Rating Plan," *PCAS*, XXIX (1942), 96.
12. WOODDY, JOHN C. *Risk Theory.* A monograph published privately in 1963 by the Society of Actuaries as part of its examinations' syllabus.

DISCUSSION OF PRECEDING PAPER

ROBERT C. TOOKEY:

Hats off to Mr. Bartlett for writing a splendid paper! The logical order of presentation made for continuity, and his slick substitutions and tricky transformations should provide a stimulating experience to most readers. Of further comfort was the fact that we were treated to functions and notations considerably more familiar to actuaries in this country than those found in the European actuarial publications. He has indeed come up with an approach that circumvents the laborious calculations inherent with some of the awesome European techniques which often involve working with slowly converging series and endlessly taking numerical results out to many places to the right of the decimal point.

The author's analysis of the errors indicates that the use of the gamma distribution would usually tend to overstate the excess ratio. While this is laudable from the standpoint of conservatism, its use in computing "excess-risk" charges in the highly competitive field of group insurance might be contraindicated if more accurate results could be obtained from a different approach. In his conclusion the author mentions the limitations "inherent in the rarefied atmosphere of the mathematical analysis," an observation I rather plaintively made several years ago. Therefore, I would like to propose a "voyage to the center of the earth" and consideration of perhaps the most brute force approach presently available for determination of excess-risk measurement in specific cases—namely, the simulation of experience through Monte Carlo techniques.

Although the author suggests this Monte Carlo approach is impractical on a case-by-case basis for small groups, my own limited experience in this area indicates otherwise. As actuaries become more familiar with the short cuts possible, the computer time for a given Monte Carlo project, say, simulation of one thousand years of experience, can be greatly reduced. The use of Monte Carlo techniques on an almost case-by-case basis for groups of over one hundred lives might be made economically feasible through the use of (1) ingenious programming, (2) adoption of approximation techniques that eliminate taking account of situations of relatively infinitesimal likelihood which thus greatly reduces the machine time required, and (3) the rounding-off of basic probability functions (e.g., mortality rate q_x and disability rate r_x) to as few significant digits as possible, in recognition of the fact that we are dealing with approximations rather than exactitudes. Step (3) allows much greater use to be made of pseudorandom numbers as they are generated, since it minimizes the number of digits required per simulated life year of exposure. The savings in computer time is directly proportional to this reduction in required digits. In the group disability indemnity situation, r_x might be taken to the nearest one hundredth. If the early durations of disability were handled as in Mr. Bartlett's example (under his thirteen-week plan) with durations following the twenty-sixth week of disability recorded to the nearest two weeks, the experience under a plan offering a maximum of two years of benefits could be simulated as soon as the exposure characteristics of the group (age, sex, and amount of weekly indemnity) had been recorded on cards or tape. Not more than five minutes of IBM 1620 computer time should be needed, since only four random digits would be required for each life year of exposure (two for r_x and two for duration). In the case of a one hundred life group life policy, the death rates would be rounded to the nearest one thousandth. One thousand exposure years for the entire group might be simulated in about one minute.

Using the frequency distribution resulting from the foregoing simulated experience, the "extra-risk" charge as a percentage of expected claims could be computed by approximate integration. These results could then be compared to those produced by Mr. Bartlett's approach of utilizing a gamma distribution that incorporates the first and the second moments of the secondary distribution. Very possibly these results could be reconciled, and an appropriate mathematical model (confirmed by the Monte Carlo tests), with an available set of tables, could be utilized in determining excess-risk charges for various groups classified in accordance with specific characteristics.

DONALD A. JONES:

The purpose of this discussion is to relate Mr. Bartlett's ideas to those of Robert H. Taylor as published in *The Proceedings of the Conference of Actuaries in Public Practice*, II (1952), 100–150. Roughly speaking, both men approximated the probability distribution of an aggregate loss random variable, say, Z, by use of a gamma distribution.

BARTLETT	TAYLOR
Aggregate Loss Random Variable	
Z = Total claims under a group insurance contract in a fixed time period.	Z = Total discounted value of future payments due to a fixed number of life annuitants.

Distribution Theory Axioms

Collective risk model	Individual risk model
Moment Formulas	
$\mu_1 = \lambda \mu_1^A$	$\mu_1 = \sum_1^N \mu_1(j)$
$\mu_2 = \lambda \mu_2'^A$	$\mu_2 = \sum_1^N \mu_2(j)$
$\mu_3 = \lambda \mu_3'^A$	$\mu_3 = \sum_1^N \mu_3(j)$
λ is the expected number of claims and $\mu_i'^A$ is the ith moment about the *origin* for the secondary distribution	N is the number of annuitants and $\mu_i(j)$ is the ith central moment of the distribution of the discounted value of future payments due the jth annuitants. (See W. O. Menge, *R.A.I.A.*, XXVI [1937], 65–88.)
Objective	
$\phi(n) = \int_n^\infty (z-n) f(z)\, dz$	$\int_n^\infty f(z)\, dz$
Approximation Method	
Fit a gamma distribution by equating first and second moments.	Fit a "translated gamma" distribution by equating first, second, and third moments.

A translated gamma distribution (abb. T-gamma in what follows) is one with the density function:

$$g(z: \alpha, \beta, \Delta) = \frac{1}{\alpha!\beta^{\alpha+1}} (z-\Delta)^\alpha e^{-(z-\Delta)/\beta}; \qquad \alpha > -1, \quad \beta > 0, \quad z \geq \Delta$$

that is, a gamma distribution translated Δ units to the right.

Taylor's approximation by fitting a T-gamma distribution has three advantages: (i) the T-gamma family has three parameters, hence "closer fits" should prevail; (ii) if the fit is by equating moments, the third moment is exact; and (iii) it is consistent with Bartlett's fit of a gamma distribution in the following sense: if μ_1, μ_2, and μ_3 satisfy the condition of gamma moments, that is,

$$2\mu_2^2 = \mu_1\mu_3 ,$$

then the fitted T-gamma will coincide with the gamma distribution fitted by equating the first moments and the second moments.

Taylor's approximation has two disadvantages: (i) μ_3 must be calculated and (ii) if $\Delta > 0$, then the probability of aggregate claims being in the interval from 0 to Δ is zero; if $\Delta < 0$, then the probability of aggregate claims being in the interval from Δ to 0 (hence negative) is positive.

Since the interesting values of the excess ratio function do not depend upon the details of the left tail of the distribution of the aggregate loss variable, this second disadvantage is not serious, and hence Taylor's T-gamma distributions would seem to deserve consideration in the calculation of the excess ratio function.

Let $\phi_T(n)$ be the excess ratio function as defined by a T-gamma distribution, that is,

$$\phi_T(n) = \int_n^\infty (z-n)\, g(z\colon \alpha_T,\ \beta_T, \Delta)\, dz,$$

where

$$\begin{aligned} \beta_T(\alpha_T+1)+\Delta &= \mu_1 & \qquad & & \beta_T &= \mu_3/2\mu_2 \\ \beta_T^2(\alpha_T+1) &= \mu_2 & & \text{or} & \alpha_T+1 &= 4\mu_2^3/\mu_3^2 \\ 2\beta_T^3(\alpha_T+1) &= \mu_3 & & & \Delta &= \mu_1 - (2\mu_2^2/\mu_3)\,. \end{aligned}$$

Substitute $x = z - \Delta$ in the integral to obtain

$$\begin{aligned} \phi_T(n) &= \int_{n-\Delta}^\infty [x-(n-\Delta)]\, g(x+\Delta\colon \alpha_T,\ \beta_T,\ \Delta)\, dx \qquad (1) \\ &= \phi(n-\Delta) \end{aligned}$$

for the α_T, β_T gamma distribution. This equation indicates that the only extra calculation required to use a T-gamma distribution is the calculation of μ_3 and Δ.

We may use equation (1) to find $\phi_T(n)$ from Bartlett's Table 1 as follows. First, putting parameter values in place of the moments we have

$$\phi(n)\frac{\mu_1}{\mu_2} = \phi(n)\,\frac{1}{\beta}$$

$$X = \frac{(\mu_1)^2}{\mu_2} = \alpha + 1$$

$$Y = \frac{n}{\mu_1}\,X = \frac{n}{\beta}.$$

Since we want $\phi(n-\Delta)$ for α_T and β_T, we enter Table 1 at

$$X = \alpha_T + 1 = 4\mu_2^3/\mu_3^2$$

and

$$\frac{Y}{X}=\frac{n-\Delta}{\beta_T(a_T+1)}=\frac{\mu_3(n-\mu_1)}{2\mu_2^2}+1$$

to read

$$\phi(n-\Delta)\,\frac{1}{\beta_T}=\phi_T(n)\,\frac{2\mu_2}{\mu_3}.$$

These formulas were applied to Bartlett's example in Section VII.

$$\mu_1 = (14.63)\,(31.353) = 458.6944$$

$$\mu_2 = (14.63)\,(1861.70)$$

$$\mu_3 = (14.63)\,(139,\,531.07571)$$

$$X = 19.387$$

$$Y/X = 1.12622$$

$\phi_T(1.2\mu_1) = 31.98$ by linear interpolation in Table 1. One may also write

$$\phi_T(n)=\sqrt{\frac{X}{\mu_2}}\left[S\left(\frac{Z}{\sqrt{X}}:\frac{2}{\sqrt{X}}\right)-S\left(\frac{Z-1}{\sqrt{X+1}}:\frac{2}{\sqrt{X+1}}\right)\right]$$
$$-Z\left[1-S\left(\frac{Z}{\sqrt{X}}:\frac{2}{\sqrt{X}}\right)\right],$$

where $Z = Y - X$ and $S(t{:}\delta)$ is the distribution function for a standardized gamma distribution with skewness δ. The function S was tabled by L. R. Salvosa in *Annals of Mathematical Statistics* (1930), 191–98 and Appendix (pp. 1–187). By use of Salvosa's tables and this formula, $\phi_T(1.2\mu_1) = 31.98$ is also obtained.

Bartlett obtained $\phi(1.2\mu_1) = 32.24$. One would expect this larger value, since Bartlett's directly fitted gamma distribution overstates the skewness.

One point, not related to Taylor's work, might be worth mentioning. Each value in Table 2 is the sum of the corresponding value in Table 1, and $Y - X$, that is,

$$\sigma(n)\,\frac{\mu_1}{\mu_2}=\phi(n)\,\frac{\mu_1}{\mu_2}+Y-X.$$

This may be verified as follows:

$$n-\mu_1=\int_{-\infty}^{\infty}(n-x)f(x)\,dx=\sigma(n)-\phi(n).$$

Now multiply each side by μ_1/μ_2.

PAUL H. JACKSON:

Mr. Bartlett's paper is a valuable addition indeed to the papers previously published in the *Transactions* on this subject. In general, the methods suggested appear eminently practical for most group insurance problems. The paper sparkles with classical statistical techniques, and the author also covers, albeit in capsule form, many of the practical limitations contained in my own written discussion of Kahn's paper.

The practical problem in experience rating is to determine an appropriate function $f(x)$ which closely approximates the frequency distribution of observed loss ratios. Both Bartlett and Kahn break this problem down into two pieces by considering the frequency distribution of loss ratios to be a compound distribution (i.e., a combination of the frequency distribution of the number of claims and the frequency distribution of claim amount). While this approach seems artificial, it has been found useful by European actuaries in their treatment of risk problems for various casualty coverages, notably fire insurance. Given that an event (a fire) has occurred, the amount of loss is variable even when one possesses the further information as to the value of the structure or, going still further, the particular structure involved. Group medical expense and disability income claims have this same characteristic, but group life claims do not. If one's knowledge is limited to the fact that exactly one death has occurred, the frequency distribution for claim amount has inherently different characteristics than for medical expense coverages (except where the insurance schedule is based on some multiple of annual earnings rounded to the nearest dollar) because only a small number of separate claim amounts are possible. Further, when the identity of the claimant is given, the amount of loss is determined completely. This qualitative difference permits more powerful statistical methods such as those suggested by Feay. With the development of survivors' income plans and more extensive group life disability benefits, this distinction, however, becomes less important.

Treating $f(x)$ as a compound distribution doubles the curve-fitting problem, but both Kahn and Bartlett as well as most other authorities agree on the Poisson distribution for the number of claims. This distribution is valid only where the probability of claim is small and where the independence of the events leading to claim can be assumed. Further, the Poisson distribution is a theoretical one which has no upper bound as to the number of possible claims, whereas in most practical applications an upper bound does exist.

I cannot agree that the assumption of independence of the events leading to individual claims rules out only epidemics, wars, and natural catastrophes. No matter how small the group, the multiple exposure resulting from air travel, commuting to and from work by train or car, working in a common location, etc., is significant in relation to the total exposure. Thus, I have more trouble with Mr. Bartlett's first condition for a Poisson process than with his second.

General considerations, as well as intuition, have led me to the conclusion that $f(x)$ can be closely approximated by a Pearson Class III curve in most practical applications. In particular, the incomplete gamma function seems eminently suitable for group insurance applications. By arriving at the gamma variate directly, without assuming the Poisson distribution for number of claims, I am led to a different interpretation of the errors discussed in Part II of Mr. Bartlett's paper.

The compound Poisson process can be accepted as a practical tool which enables us to determine the parameters for the gamma distribution from the moments of the secondary distribution of claim amount. The skewness of the resulting gamma distribution is larger than that of the particular compound Poisson distribution and has a larger area under the right-hand tail which produces larger values for the excess ratio function. The differences do not, for me at least, represent the degree of overstatement of excess ratio computed from the gamma distribution over the "true" ratio computed from the corresponding compound Poisson distribution, but rather the degree of understatement of excess ratio computed from the compound Poisson distribution as compared with the "true" ratio computed from the gamma distribution.

This slightly different approach has two important implications. First, it means that the actuary employing Mr. Bartlett's method will not view the "errors" and "overstatements" as an indirect loading for conservatism. Second, in the case of a group life plan with uniform amounts, the actuary would still use the gamma variate rather than the simple Poisson distribution. It can be shown, for example, using Mr. Bartlett's method, that in this instance the secondary distribution is a "spike" distribution and the gamma variate has parameters $\beta = 1$ and $\alpha = \lambda - 1$ so that $f(x)$ reduces to

$$\frac{X^{\lambda-1}e^{-x}}{(\lambda-1)!}$$

rather than

$$\frac{\lambda^{x}e^{-\lambda}}{X!},$$

the simple Poisson distribution.

It might be noted that the ratio of this gamma probability density to Poisson is less than unity near the mean and greater than unity for sufficiently large X, since the ratio is dominated by $X!$ as X increases without bound. Where λ is an integer greater than one and $X = \lambda + 1$, the ratio is always less than unity, since $(n + 1)^n \cdot n^{-n}$ approaches a limit with increasing n that is less than or equal to e. At any rate, if the compound Poisson distribution is considered to be the true underlying frequency distribution, then the actuary must use the simple Poisson distribution for cases with uniform amounts and shift gears to gamma when the first minor difference in insurance amounts appears, and this discontinuity can be avoided by simple shift of the "true" label over to the gamma variate.

In group life insurance applications, insurance amounts are usually related to salary, thus suggesting possible use of the logarithmico-normal distribution as the secondary distribution for claim amount. For large amounts of coverage which are related to higher incomes and which are frequently "pooled" on a company-wide basis, the Pareto distribution appears appropriate (see K. C. Hagstroem, "Inkomstatjamningen i Sverige," *Skand. Bankens Kvart.-Skr.*, April, 1944).

For stop-loss computations at high limits, the assumption of claim independence and normal distribution of claim amount appear unjustifiable. The entire process of fitting $f(x)$ to actual data involves normal data and the subsequent use of $f(x)$ only in the area of abnormal losses is best characterized as "non-Bayesian." The group writing companies are understandably disinterested in offering true stop-loss insurance for self-insured plans, and it is likely that there will never be sufficient data accumulated to enable the actuary to rely solely on classical statistical techniques to determine high-limit stop-loss premiums. In fact, it appears that the common practice of group writing companies is to restrict coverage in such a way that they limit their bet to the sure thing and, thus, have no exposure at all in this area. This is unfortunate since their knowledge in this area and their ability to provide true catastrophe insurance is perhaps their greatest defense against the further encroachment of self-insurance.

Mr. Bartlett's masterly work has eminently practical applications for experience rating and I hope that the reservations expressed herein in no way imply anything short of profound admiration on my part.

(AUTHOR'S REVIEW OF DISCUSSION)

DWIGHT K. BARTLETT, III:

I would like to thank Messrs. Tookey, Jones, and Jackson. They do me a great honor by their thoughtful discussions of my paper.

Mr. Tookey's plea for a "voyage to the center of the earth" is a persuasive one. It is worthwhile for actuaries concerned with experience rating problems but who have a limited exposure to computers to know that Monte Carlo simulations of mortality and morbidity experience on insured groups can be done so quickly and efficiently. Perhaps someone will feel moved to include in our literature as an actuarial note the Fortran based on the most efficient techniques of the Monte Carlo simulation. However, actuaries who do not have access to large-scale computers will still have to rely on classical statistical techniques such as those used in my paper.

Mr. Jones's discussion bringing the ideas of Mr. Taylor to our attention is particularly valuable. Mr. Taylor's method of fitting a "translated gamma" distribution appears to be a significant improvement over that proposed in my paper. While the practical effect was very small in the example included in Section VII of my paper, it would perhaps be much more significant in examples with a smaller claim frequency or a higher stop-loss claim level.

Mr. Jackson has stressed the weaknesses in the assumptions of the Poisson process as applied to claims on insured groups. He particularly brings out the exposure to multiple-claim situations involved in mass transportation, common working location, etc. This same criticism can be made, of course, of the Monte Carlo technique, which also assumes independence of the probability of occurrence of individual claims. It is not clear exactly how great a weakness this is. It intuitively seems to me to be a greater weakness in life insurance coverages than in health insurance. Perhaps companies that have been writing group accidental death catastrophe reinsurance can contribute to the literature their experience and their methods of premium rating which will shed some light to the profession in general on this problem.

After stating his position in favor of using the gamma distribution rather than the compound Poisson distribution as being the "true" distribution, Mr. Jackson then apparently resorts to the compound Poisson distribution as a basis for determining the parameters for the gamma distribution, just as I have done in the paper. This would appear to be a little like wanting your cake and eating it too.

Mr. Jackson's comments about the techniques of the paper being "non-Bayesian" are certainly correct. I hesitate to comment further on this point in view of my lack of education in this newly developing field of statistics. However, while I stand ready to be corrected on this point, let me say that it would seem to me that the more strongly a Bayesian feels about the prior distribution the more confidence he will have in his posterior distribution. If this is true, the techniques of the paper might still be of value to the Bayesian by giving him a prior distribution in which he might have a fairly strong degree of confidence.