AN EXPERIENCE RATING FORMULA (REPRINT)

Ralph Keffer

An Experience Rating Formula
by
Ralph Keffer

This short paper, published in 1929, is reprinted, by permission, from the *Transactions of the Actuarial Society of America*, which was the predecessor of the Society of Actuaries. It is the earliest known application of the gamma-Poisson mixture to experience rating, and seems remarkably modern. Note also three references to PCAS papers, which suggests there was a fair degree of interaction between life and casualty actuaries of that era.

An Experience Rating Formula

BY

RALPH KEFFER.

Mr. Albert W. Whitney has developed a formula for experience rating which is described in a paper appearing in Volume IV of the Proceedings of the Casualty Actuarial Society. This formula was developed from the standpoint of Workmen's Compensation Insurance, but it has been adapted to other lines, in particular to Group Insurance.*

Mr. Whitney assumed that, for any given class of risks, the average class hazard resulted from different individual risk hazards. In order to develop a formula he assumed that these individual risk hazards were distributed about the mean class hazard in accordance with a known frequency curve. For the purpose of his paper he assumed that the normal frequency curve would apply. Then, on the assumption of this frequency distribution of the real risk hazard, the problem which he set was to develop a formula for the most probable rate which, when applied to a particular individual risk, would make possible the actual experience which was observed. The formula developed on this assumption did not appear to be workable from a practical standpoint and therefore various substitutions and approximations have been suggested for the term z which appears in the formula

$$x = P + z\,(p - P)$$

but the form which seemed to be preferred was

$$z = \frac{Pn}{Pn + K}$$

where Pn is the total premium for the risk and K is a constant to be determined by judgment and inspection.

In the consideration of some questions relating to Group Insurance my attention was called to a certain formula which proved to be Mr. Whitney's formula in a little different form. This led to the investigation of the assumptions underlying the formula

* See, for example, remarks by Mr. Bassford, P.C.A.S., Vol. VIII, p. 307.

with particular reference to the meaning of the constant in the formula. A development of this formula is given below, starting from certain original assumptions which differ somewhat from those made by Mr. Whitney. The formula has been considered with particular reference to its application to Group Insurance although it would apply in certain other lines of insurance.

The following are the initial assumptions:

(1) Assume the existence of an average scale of net rates of mortality which when applied to all groups or to all groups of a certain classification will give the real expected number of deaths for the combined groups.

(2) Assume the existence of a true scale of rates of mortality for any individual group such that the variations in actual experience from year to year from this true rate are in accordance with the laws of probability.

(3) Assume this true scale of rates for each individual group may be obtained by multiplying the rates for each age of the average scale by a constant.

(4) Assume the average scale of rates for all groups combined does not change during the period under observation.

(5) Assume the true scale of rates for an individual group does not change during the period of observation.

(6) Assume the ratios of the true scale of rates for each group to the average scale for all groups combined are distributed about the mean in accordance with the following frequency distribution:*

$$y = Ce^{-kr}(kr)^m \tag{1}$$

where r is the ratio of the true rate to the average rate and C, k, and m are constants to be determined.

This frequency distribution appears more natural to use than the normal since $y = o$ for $r = o$ and y has a finite value for

* This is a special form of Pearson's Type III frequency curve. See Elderton "Frequency Curves and Correlation." The equation there is in the form

$$y = y_0 e^{-\gamma x}\left(1 + \frac{x}{a}\right)^{\gamma a}$$

but it may be changed to the form of equation (1) by taking $\gamma = 1$ and making the substitutions $a = m$ and $a + x = kr$ after which $C = \frac{y_0 e^m}{m^m}$.

every value of r greater than zero. The ratio of the true scale of rates to the average scale must be greater than zero, but there is not necessarily an upper limit to its value.

The following considerations determine the values of C and k. If the equation is to be expressed in a form such that

$\int_{r_1}^{r_2} y\,dr$ will give the probability that r lies between r_1 and r_2

then the constant C must be determined so that $\int_0^\infty y\,dr = 1$.

But $\int_0^\infty C\,e^{-kr}\,(kr)^m\,dr = \frac{C}{k}\,m!$*

$$\therefore C = \frac{k}{m!}.$$

By definition the mean value of r is 1. But the mean value of r is given by

$$\frac{\int_0^\infty r\,y\,dr}{\int_0^\infty y\,dr} = \int_0^\infty \frac{kr\,e^{-kr}\,(kr)^m}{m!}\,dr$$

$$= \frac{m+1}{k}\int_0^\infty \frac{e^{-kr}\,(kr)^{m+1}}{(m+1)!}\,d\,(kr)$$

$$= \frac{m+1}{k}$$

$$\therefore k = m + 1.$$

The equation of the frequency curve is reduced to the form

$$y = \frac{(m+1)\,e^{-(m+1)r}\,[(m+1)r]^m}{m!} \tag{2}$$

which contains the as yet undetermined constant m.

To see the effect of the constant m in equation (2) it may be simpler to make the substitution

$$x = (m+1)r$$

after which equation (2) reduces to the form

$$y = \frac{(m+1)\,e^{-x}\,x^m}{m!}. \tag{3}$$

* This integral is a form of the Gamma function

$\Gamma(n+1) = \int_0^\infty e^{-x}\,x^n\,dx = n!$ for integral values of n.

The constant m determines the shape of the curve and hence depends on the assumptions regarding the distribution of all possible true rates of mortality. A large value of m means that they are assumed to be closely grouped about the mean, i.e., that the *a priori* probability that the true rate is near the average rate is very high.

The graphs of equation (2) for values of $m = 14$, 29, 44 and 89, show the effect of different values of m.

The total area under each curve is unity and the area under any curve between any two limits is equal to the assumed *a priori* probability that the true rate applicable to a group about which nothing is known, will lie between those limits.

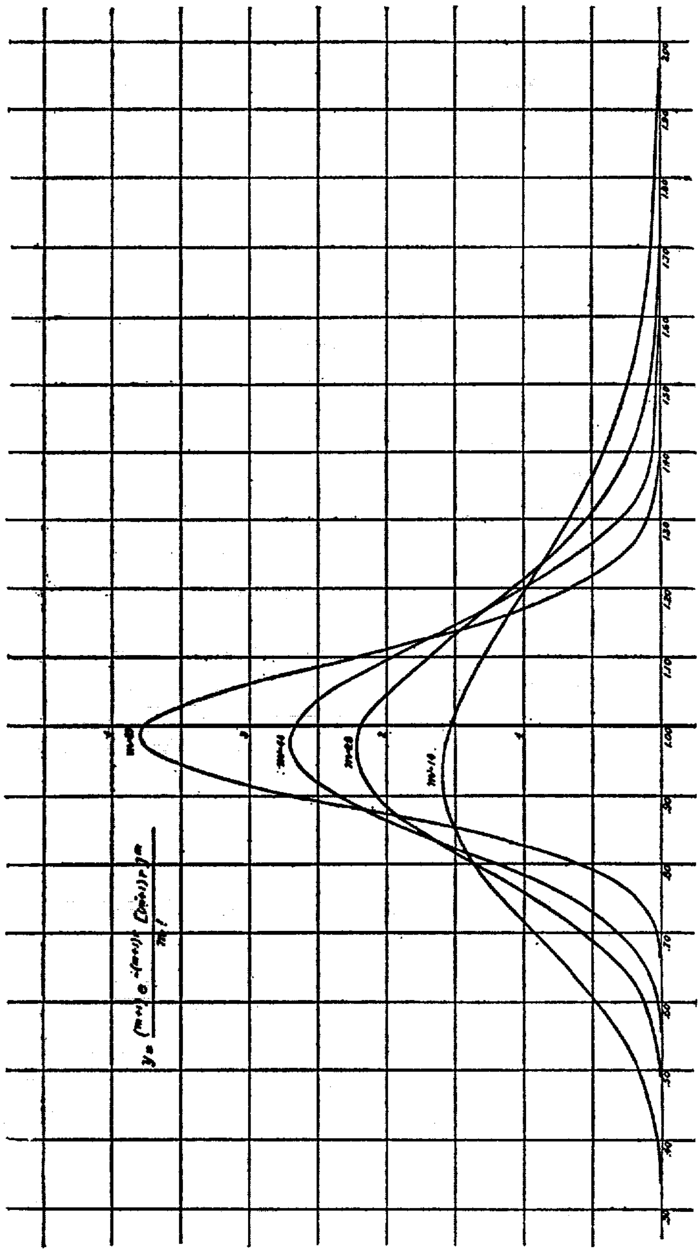
The following table summarizes the values of these probabilities:

ASSUMED DISTRIBUTION OF TRUE RATES OF MORTALITY.

Ratio of True Rate to Average Rate	Percentage of total groups which may be expected to fall in each class			
	$m = 14$	$m = 29$	$m = 44$	$m = 89$
30% to 40%	.1%			
40% to 50%	.9%			
50% to 60%	3.1%	.5%	.1%	
60% to 70%	6.9%	3.1%	1.3%	.1%
70% to 80%	11.6%	9.6%	6.9%	2.3%
80% to 90%	14.9%	17.6%	18.2%	14.9%
90% to 100%	15.9%	21.9%	26.4%	34.7%
100% to 110%	14.4%	19.9%	23.7%	31.0%
110% to 120%	11.5%	13.7%	14.9%	13.5%
120% to 130%	8.3%	7.9%	6.3%	3.2%
130% to 140%	5.4%	3.7%	1.5%	.3%
140% to 150%	3.4%	1.5%	.5%	
150% to 160%	1.9%	.5%	.2%	
160% to 170%	1.0%	.1%		
170% to 180%	.5%			
180% to 190%	.2%			

The use of $m = 89$ implies that the true rates of mortality will be practically confined between the limits of 70% and 130% of the average rate with 94% of the cases between 80% and 120% while the use of $m = 14$ implies a wider spread from 40% to 180% of the average rate with only 57% of the cases between 80% and 120%. At the present time there does not seem to be any way to fix a value of m except to estimate the probable range by judgment.

In each case the mean value of r is at the point $r = 1$, but the



mean value is not the most probable value as may be seen from the curves or as may be determined analytically by setting the first derivative of y equal to zero. This shows the most probable value to be at the point

$$x = m$$

$$\text{or } r = \frac{m}{m+1}.$$

Moreover, the probability that r is less than 1 is greater than the probability that r is greater than 1, which means that if the true rate could be determined for each group a larger number of groups would be entitled to reductions below the average rate than would require increases. This is to be expected because of two groups of the same size, the one with the greater number of expected deaths will contribute more to the average experience. For any given group r is equal to the ratio of the expected deaths at the true rate to the expected at the average rate.

Let d be the actual number of deaths in a given group over a period of time for which the expected number at the average rate is c.

Then rc is the expected number of deaths at the true rate.

Since the probability of death is small, we may assume that Poisson's formula* holds for the probability that a given number of deaths will occur, therefore the probability that d deaths will result when the true expected is rc is

$$\frac{e^{-rc}\,(rc)^d}{d\,!}.$$

But from our assumed frequency distribution the probability that the true r lies between r and $r + dr$ is

$$\frac{(m+1)\;e^{-(m+1)r}\;[(m+1)r]^m}{m\,!}\,dr.$$

Therefore, the probability that the true rate r lies between r and $r + dr$ and that the application of this rate r to a given group in which the expected number of deaths at the average rate is c,

* Sometimes known as the Bortkewitsch "Law of Small Numbers." See description of Table LI in Pearson's Tables for Statisticians and Biometricians or Fisher, Mathematical Theory of Probabilities, 2nd Edition, p. 265, etc.

will result in d deaths is the product of the above two expressions, which may be put into the following form

$$\frac{(m+1)^{m+1}c^d(m+d)!}{(m+1+c)^{m+d}\,m!\,d!}\cdot\frac{e^{-(m+1+c)r}[(m+1+c)r]^{m+d}}{(m+d)!}dr.$$

Hence the distribution of the values of r which will result in d actual deaths in groups for which the expected at average rate is c, is given by the curve

$$y=K\frac{e^{-(m+1+c)r}[(m+1+c)r]^{m+d}}{(m+d)!} \tag{4}$$

where K is the constant multiplier which appears in the previous expression.

By an analysis similar to that used for equation (1), the mean value of $(m+1+c)r$ is found to be at the point

$$(m+1+c)r=m+d+1$$

hence the mean value of the ratio of the true rate to the average rate for a group where the actual number of deaths is d and the expected at average rate is c is

$$r=\frac{m+1+d}{m+1+c}. \tag{5}$$

In order to compare with Mr. Whitney's formula this may be written as

$$r=1+\frac{c}{c+m+1}\left(\frac{d}{c}-1\right). \tag{6}$$

By differentiating the expression in equation (4) we find that the most probable value of the ratio of the true rate to the average rate for a group where the actual number of deaths is d and the expected at the average rate is c, is

$$r=\frac{m+d}{m+1+c}.$$

But the most probable value is not necessarily very probable and for insurance purposes the mean is the more logical function to use. In this particular case there is little difference between the mean and the most probable unless small values of m are assumed.

The c in formula (5) is the expected number of deaths at the average rate applicable to all groups or to all groups of a certain class. The total group experience of six companies has been compiled each year and ratios of actual to expected by the

American Men Ultimate Table have been published.* This experience is large enough to give accurate results not only for the total experience but for certain subdivisions.

This experience is easily converted into loss ratios in terms of the Standard Gross Premium Rate prescribed by the State of New York. The formula used to compute this scale of rates is

$$P_{\overset{1}{x}\,\overline{1|}} = \frac{(1.035)^{-1}\, q_x + .0017}{.935}$$

where q_x is by the $AM^{(5)}$ table.

The total premium is then

$$\Sigma\, P_{\overset{1}{x}\,\overline{1|}} = \frac{(1.035)^{-1}\, \Sigma\, q_x + .0017\, S}{.935}$$

where S is the total amount of insurance exposed and $\Sigma\, q_x$ is the total expected mortality by the $AM^{(5)}$ table, both of which values are given in the tabulations of the group experience.

The ratio of actual claims to total premiums will give the loss ratio at standard rates. Let this average loss ratio for all groups combined be λ. Then, instead of r in formula (5) we shall want to find $\lambda\, r$ to determine the portion of the premium at standard rates that we shall require for payment of claims. The c in formula (5) may be expressed in terms of loss ratios at standard rates and the formula transformed in several ways for ease of computation.

Formula (5) is expressed in terms of number of deaths and this is essential to its theoretical development. For practical purposes it may be expressed in terms of amounts of insurance on the assumption that the experience will be the same as if each life were insured for the average amount. The formula on this basis may be written as follows

$$r = \frac{(m+1)\, A + D}{(m+1)\, A + C} \qquad (7)$$

where A is the average amount of insurance in force upon each life in the group and D and C are respectively the actual and expected losses by amounts. If a death loss occurs for an amount in excess of the average, formula (7) would then give

* T. A. S. A., Vol. XXVI, p. 332, also privately published annual reports by E. E. Cammack, Chairman of the Committee on Group Mortality Investigations.

a higher value of r than if formula (5) were applied. Since abnormal losses are more likely to be in excess of the average than otherwise, this modification of the formula would, in general, be on the safe side.

In the practical use of formula (5) or (7) allowance must be made for incurred and unreported claims. This may be done by making a deduction from c or C or by deferring the application of the formula to a given experience until all claims are likely to be reported.

It must be kept clearly in mind that formula (5) does not necessarily give an approximation to the true rate for any given group. The ultimate experience for any given group may be found to be different from either the average rate for all groups or the first rate given by the use of the formula. However, as the experience increases and c and d become large in comparison with m the formula gives a rate which is nearer the indicated rate. From the probability theory we know that the indicated rate will approach the true rate as experience increases so that for large enough groups the formula should give a satisfactory approximation to the true rate.

What the formula does is to give a reclassification of the groups by size and experience. It determines a new average rate for each new class such that for a large enough business the premium income should, in the aggregate, be the same as if the uniform average rate were charged each group. If we have two groups of the same size with the same number of deaths in the past, the true mortality rate of one may be quite different from the true mortality rate of the other; nevertheless, in the absence of other information bearing on the risk it seems proper that the same premium rate should be made applicable to each. From this point of view formula (5) may be said to determine the best rate of mortality to apply to a given group, subject of course, to the original assumptions of this paper being applicable to the group business.

The question of experience rating for group insurance must be considered by two types of companies. On the one hand there is the non-participating company which expects to charge a uniform average rate for all groups the first policy year, but expects to adjust future rates on the basis of experience. For such companies the formula offers a satisfactory method of determining future rates. If a uniform rate is charged the first year which produces a total premium just sufficient to pay claims and overhead, then upon renewal it will be necessary to make increases for some groups if reductions are made for others.

On the other hand a company which issues participating group policies expects to charge each year a premium which will be greater than required and then expects to adjust the net cost by dividends at the end of the year. Formula (5) may be used by such companies to determine the portion of the premium paid by each group which should be applied to mortality, for the formula applies just as well to past experience as to future. Logically it seems to be the proper basis of a method of distributing dividends, for it determines the rate that would have been charged at the beginning of the year if there had then been available the knowledge regarding the risk which developed during the year. But practical questions enter into its adaptation to distribution of dividends unless the premium rate charged contains a sufficient margin to cover the mortality of the most unfavorable group. The question of negative dividends brought about when the participating premium is insufficient has been considered in a paper by Mr. William Leslie* to which reference should be made. As pointed out there and in the discussion by Mr. Bassford, retroactive increases in premium are usually uncollectable and so if the original gross premium is not sufficient to cover the adverse mortality in certain risks, the deficit must be made up elsewhere. All dividends may be reduced or a maximum dividend rate may be adopted in which case the groups with good experience will not receive the full dividends to which they would otherwise be entitled or an increase in future rates may be counted upon to make up past deficits as well as to provide an adequate rate for the future.

The practical application of any experience rating formula or dividend distribution formula must, of course, take account of expenses, but this paper has been limited to a consideration of the mortality factor alone.

* P. C. A. S., Vol. VIII, pp. 70-71. See also discussion, pp. 308-309.