

**ON THE CREDIBILITY OF GROUP INSURANCE  
CLAIM EXPERIENCE**

**MYRON H. MARGOLIN**

**ABSTRACT**

This paper deals with credibility as applied to group insurance, where the problem is to estimate the future claim rate of a group as a function of both its own past claim rate and the average claim rate observed for similar groups.

First, the concept of the manual premium rate is briefly discussed. The manual rate embodies the actuary's subjective estimate of future claim experience and has only limited basis in fact.

The approach to credibility is to measure the correlation among successive years' claim rates, where claim rates are expressed in relation to those rates anticipated by the manual premium rating system. Credibility factors are determined as the correlation coefficients in a multiple regression equation relating one year's claim rate to the claim rates in prior years and to the average claim rate for all similar groups.

A sample of the application of the new theory to some actual comprehensive major medical claim experience discloses that, for this coverage, one year of prior claim experience is almost as good an indicator of future experience as two years.

The paper concludes with a brief comparison with previous theories of credibility. Previous theories have generally rested upon the hypothetical existence of "true" claim rates underlying the rates actually observed, but there is no reason to believe in their existence.

---

**I. INTRODUCTION: THE CONCEPT OF THE MANUAL RATE**

**T**HE basic problem with which credibility theory deals is that of estimating the future claim experience of a group as a function of both its own actual past experience and the average experience expected for all groups of the same type. It is generally assumed that for small groups the actual past experience deserves very little credibility in relation to the average, whereas for very large groups it may deserve full or nearly full credibility, and that for intermediate sizes some blend of actual and average experience should apply.

In reality, the phrase "average experience expected for all group cases of the same type" lacks a clear definition. No two groups are exactly alike, and some are strikingly and obviously different. Moreover, group benefit plans are frequently unique. How can one strike an average for what is really a heterogeneous lot? The answer is to construct a manual premium rating system.

"Manual premium" is the term commonly used in group insurance to refer to one's a priori estimate of future claims, plus loadings for expenses and contingencies. Apart from variations in the loadings for different size cases, the manual rate (i.e., manual premium per unit of exposure) is the rate one charges in the absence of any knowledge of actual claim experience for a group and is in direct proportion to the expected claims. Based on knowledge of the claim experience of other groups, the manual rate reflects how such factors as the plan details, age distribution, female percentage, and geographical area are expected to affect the claim experience of a new group.

Obviously, no company can engage in a group insurance operation without some sort of manual rating system, unless (a) it underwrites only groups which have prior claim experience and (b) it gives full credibility to this experience. Generally each company has its own system. Some of these systems may be better than others, but in practice none is perfect—that is, able to predict exactly the future claim experience of most groups. It would be rather pointless for actuaries to try to decide, once and for all, on a priori grounds, which of two systems is superior. Each is based on only very imperfect knowledge of what has actually happened—that is, of what the true characteristics of the various groups were and how the actual claim experience related to these characteristics. The development and application of manual rating schemes entail much personal judgment. Furthermore, the actuary or group underwriter will inevitably confront new cases or situations in which he will, in a sense, rewrite the manual and invent subjective new rules to fit what he regards as unprecedented. (From a conceptual point of view, the manual rating system is analogous to Bayesian statistical techniques.<sup>1</sup> The manual rating schemes developed by the various insurance companies and group actuaries are in essence compilations of personalistic or subjective probabilities with only limited bases in fact.)

Accordingly, the balance of this paper will presuppose the use of some

<sup>1</sup> For an introduction to the subject see D. A. Jones, "Bayesian Statistics," *TSA*, XVII (1965), 33.

manual premium system, without attempting to describe exactly how such a system might be derived. Further, as is the practice in group insurance, claim rates will be expressed in relation to manual premiums, as "loss ratios" or percentages of the manual premiums produced by the system. Let us denote by  $\bar{X}$  the common a priori expected loss ratio for all cases, assuming as well that the mean actual loss ratio for all cases combined equals  $\bar{X}$ ;  $\bar{X}$  will be treated as constant over time, with the manual rates assumed to provide properly for any inflationary trends.

After a particular case has been in force for a period of time and the actual claim experience during this time has been measured, the actuary will wish to modify the rates charged. The extent to which he supplants the a priori expected rate by the observed rate is the degree of credibility he attaches to the actual experience.

## II. CREDIBILITY AS REGRESSION

The approach to credibility presented below is essentially empirical or statistical, measuring the statistical relationship of the loss ratios of one year to those of prior years and to the mean. An attempt is made to formulate the few theoretical premises as explicitly and carefully as possible.

1. In the absence of actual claim experience for a particular group, one formulates an a priori estimate of its future claim experience using a manual rating system, which itself is derived from knowledge of actual experience on other cases. The mean,  $\bar{X}$ , is the expected loss ratio—the ratio of the a priori estimated claims to the manual rate—and is the same for all groups.
2. The actual loss ratios  $X$  for all in-force groups will form some distribution around  $\bar{X}$ . The particulars of this distribution will depend both on the properties of the groups and on the particular manual rating system in use. (The characteristics of this distribution, however, are of no concern here.)
3. Because the characteristics of each in-force group change continually and the changes are only partly reflected in the manual rates, the loss ratios of a given case differ from year to year. (Of course, they change even within a year, so that the loss ratio for the entire year formed by dividing manual premiums into incurred claims is really a kind of average.) The best estimate of  $X_{n+1}$ , the loss ratio in policy year  $n + 1$ , is neither  $X_n$  nor  $\bar{X}$ . Nonetheless, postulate a general relationship between  $X_{n+1}$  and  $\bar{X}$ ,  $X_n$ ,  $X_{n-1}$ , . . . . By examining the past, we can infer a statistical relationship between  $X_n$  and  $\bar{X}$ ,  $X_{n-1}$ ,  $X_{n-2}$ , . . . , applying to a large number of groups. We assume that the two relationships are the same—that, if we find a certain function  $F$  with arguments  $(\bar{X}; X_{n-1}, X_{n-2}, \dots)$  which would have predicted  $X_n$  with a certain degree of accuracy, then  $F(\bar{X}; X_n, X_{n-1}, \dots)$  will be an equally good predictor of  $X_{n+1}$ .

(Even the above assumption cannot always be valid. The advent of Medicare probably caused a sudden though temporary reduction in credibility factors. It may also be true that the health care situation has been more "turbulent" in recent years than previously, resulting in a permanent reduction in credibility. We have no choice but to accept the assumption, however; to reject it is tantamount to rejecting the past as a guide to the future.)

4. There exists a myriad of functions  $F$  which would have predicted  $X_n$  with varying degrees of success. Among them are various linear combinations of  $\bar{X}$ ;  $X_{n-1}$ ,  $X_{n-2}$ , . . . . Although the linear combinations are not necessarily the best predictors, they offer advantages of computational convenience. The best of the linear combinations can be ascertained by application of the least-squares principle—that is, by finding the multiple regression of  $X_n$  on the arguments—and the calculation of the regression coefficients follows well-known procedures.<sup>2</sup>

The essential features of this new approach come to light when the available data are assumed to be limited to one year of actual experience. The problem is then simply to find the regression of  $X_n$  on  $\bar{X}$  and  $X_{n-1}$ . From mathematical statistics, we know in advance that the credibility factor will be the correlation coefficient between  $X_n$  and  $X_{n-1}$ , but it is perhaps helpful to spell out the derivation.

Let  $E(X_n)$  be the predicted value of  $X_n$ , where  $E(X_n) = ZX_{n-1} + (1 - Z)\bar{X}$ ; let  $e_n = X_n - X_{n-1}$ ; and let  $f(X_n)$  be the frequency distribution of  $X_n$ . Determine the credibility factor  $Z$  by minimizing

$$\begin{aligned} & \int_{-\infty}^{+\infty} [X_n - E(X_n)]^2 f(X_n) dX_n \\ &= \int_{-\infty}^{+\infty} [X_{n-1} + e_n - ZX_{n-1} - (1 - Z)\bar{X}]^2 f(X_n) dX_n. \end{aligned} \quad (1)$$

Since there is a one-to-one correspondence between the  $X_n$ 's and the  $X_{n-1}$ 's, we can as readily integrate over the  $X_{n-1}$ 's instead of over the  $X_n$ 's. Write  $f(X_{n-1})dX_{n-1}$  for  $f(X_n)dX_n$ , and then drop the subscript  $n - 1$ ; then equation (1) becomes

$$\begin{aligned} & \int_{-\infty}^{+\infty} [X + e - ZX - (1 - Z)\bar{X}]^2 f(X) dX \\ &= \int_{-\infty}^{+\infty} [(1 - Z)(X - \bar{X}) + e]^2 f(X) dX \quad (2)^3 \\ &= (1 - Z)^2 \sigma_X^2 + 2(1 - Z) \text{cov}(X, e) + \sigma_e^2. \end{aligned}$$

<sup>2</sup> See, for example, P. G. Hoel, *Introduction to Mathematical Statistics* (1st ed.; New York and London: John Wiley & Sons, Inc., 1947), pp. 100 ff.

<sup>3</sup> The function  $\text{cov}(a, b)$  as used here means

Next differentiate with respect to  $Z$  and set

$$-2(1 - Z)\sigma_X^2 - 2 \operatorname{cov}(X, e) = 0,$$

or

$$Z = 1 + \frac{\operatorname{cov}(X, e)}{\sigma_X^2} = 1 + \frac{\operatorname{cov}(X_{n-1}, e_n)}{\sigma_{X_n}^2}. \tag{3}$$

Since  $X_n = X_{n-1} + e_n$ , it then follows that

$$\int_{-\infty}^{+\infty} X_n X_{n-1} f(X_{n-1}) dX_{n-1} = \int_{-\infty}^{+\infty} X_{n-1} (X_{n-1} + e_n) f(X_{n-1}) dX_{n-1},$$

or

$$\operatorname{cov}(X_n, X_{n-1}) = \sigma_{X_{n-1}}^2 + \operatorname{cov}(X_{n-1}, e_n).$$

Similarly,

$$\sigma_{X_n}^2 = \operatorname{cov}(X_n, X_{n-1}) + \operatorname{cov}(X_n, e_n)$$

and

$$\operatorname{cov}(X_n, e_n) = \operatorname{cov}(X_{n-1}, e_n) + \sigma_{e_n}^2.$$

If  $\sigma_{X_n}^2 = \sigma_{X_{n-1}}^2$ , it further follows that

$$\operatorname{cov}(X_n, e_n) = -\operatorname{cov}(X_{n-1}, e_n) = \frac{1}{2}\sigma_{e_n}^2.$$

Now equation (3) can be written anew as

$$\begin{aligned} Z &= 1 + \frac{\operatorname{cov}(X_{n-1}, e_n)}{\sigma_{X_n}^2} = 1 - \frac{\operatorname{cov}(X_n, e_n)}{\sigma_{X_n}^2} = 1 - \frac{\sigma_{e_n}^2}{2\sigma_{X_n}^2} \\ &= \frac{\operatorname{cov}(X_n, X_{n-1})}{\sigma_{X_n}^2} = \rho, \end{aligned} \tag{4}$$

where  $\rho$  is the correlation coefficient between  $X_n$  and  $X_{n-1}$ . Equation (4) guarantees that  $Z$  will be zero or positive, provided that the correlation is nonnegative. Given that  $\bar{X}$  and  $\sigma$  are constant, 100 per cent credibility is warranted for a class of groups only if  $\rho = 1$ , that is, if the loss ratios never change. In practice, there is no evidence that there is such a class, and, further, there are no grounds to believe that with increased size  $\rho$  approaches unity as a limit.

---


$$\int_{-\infty}^{+\infty} abf(a) da.$$

If  $a$  and  $b$  are measured from their means, this is the so-called covariance;  $\sigma_X^2$  is the variance of the loss ratios  $X$  as measured from  $\bar{X}$ .

The fact of the dependence of the credibility factors upon the particular manual rating scheme is obvious. Furthermore, the factor (or factors, if the relationship among three or more years of experience is studied) applies to a class of groups as a whole. In judging how broadly or narrowly to define the classes, the actuary must compromise between including many statistical data (many cases) in each class and maintaining as nearly as possible the homogeneity of the members of each class.

### III. A SAMPLE APPLICATION

The only data needed to put the new approach to work are successive years' loss ratios for a class of group cases. If only two years of experience are used, simply calculate the correlation coefficient.

TABLE 1  
LOSS RATIOS FOR COMPREHENSIVE MAJOR MEDICAL  
RESULTS FOR PAIRS OF YEARS

LIVES	n*	1967-68			1968-69			1969-67		
		m†	s‡	r§	m	s	r	m	s	r
1-24 . . . . .	330	0.79	0.43	0.25±0.05	0.81	0.42	0.18±0.05	0.88	0.61	0.25±0.08
25-49 . . . . .	430	0.79	0.36	0.27±0.05	0.82	0.38	0.33±0.04	0.89	0.57	0.11±0.07
50-99 . . . . .	270	0.82	0.32	0.35±0.05	0.78	0.28	0.36±0.06	0.78	0.33	0.29±0.08
100-199 . . . . .	120	0.78	0.23	0.61±0.05	0.80	0.28	0.53±0.08	0.82	0.32	0.25±0.11

\* The approximate number of cases in 1967-68 or in 1968-69. In 1969-67 the number is somewhat smaller.

† The mean loss ratio for the first year shown in each pair.

‡ The standard deviation of *m*.

§ The correlation coefficient, shown with 70 per cent confidence limits.

As an example of the application of the new approach to more than two years, Table 1 presents an analysis of the loss ratios of certain groups insured by Prudential for comprehensive major medical insurance for at least two of the policy years ending in 1967, 1968, and 1969.

The groups included are those which had policy anniversaries in the first six months of each year shown. The policies may have been issued in any prior year. Cases which changed size bracket within the years indicated are excluded. All cases were originally issued with at least twenty-five employees; some subsequently shrank below twenty-five. The loss ratios equal the incurred claims divided by the manual premiums for the policy year.

Unfortunately, the data are flawed in three fairly important respects:

1. The correlations are derived from those cases which renewed in the two successive years shown and necessarily exclude cases which canceled before the second year. This certainly introduces some sort of bias into the results.

2. The manual rate bases in use were slightly different in each of the three years. Furthermore, the manual premium for a particular case in a particular year is frequently an approximation estimated from the manual rate calculated in a previous year. The effect of this flaw is probably to reduce the correlations slightly.
3. The incurred claims from which the loss ratios were calculated are not the true incurred claims. Instead, they equal the cash claims charged to the case (i.e., paid) during the year plus the increase in the allowance for incurred but unreported claims. This allowance, which is a function of both the cash claims and the payable premium, is only an estimate of the claims which were incurred during the policy year but paid in a subsequent year.

This deficiency must serve to increase the correlation between the loss ratios for successive years; consider the effect of an illness incurred in one year which gives rise to several claim payments, some paid in one policy year and some in the next.

TABLE 2

Lives	$r_1$	$r_2$	$Z_1$	$Z_2$	$V_1$	$V_2$
1-24 . . . . .	0.25	0.16	0.22	0.10	0.938	0.928
25-49 . . . . .	0.30	0.18	0.27	0.10	0.910	0.901
50-99 . . . . .	0.40	0.23	0.37	0.08	0.840	0.834
100-199 . . . . .	0.54	0.32	0.52	0.04	0.708	0.707

Nevertheless, these flawed data are the right data to use when the object is to estimate future loss ratios emerging under similarly flawed circumstances.

From the correlation coefficients in Table 1, one can derive credibility factors applicable to  $X_{1967}$  and  $X_{1968}$  with which to estimate  $X_{1969}$ . One will first wish to graduate and otherwise adjust the raw data, however, fitting it to the two preconceptions that (a) the correlation coefficients progress smoothly by size and (b) the correlation coefficient for 1968-69 is really the same as that for 1967-68. Then<sup>4</sup>

$$E(X_{1969}) = \frac{r_1(1 - r_2)}{1 - r_1^2} X_{1968} + \frac{r_2 - r_1^2}{1 - r_1^2} X_{1967} + \frac{1 - r_2}{1 + r_1} \bar{X}, \quad (5)$$

where  $r_1$  is the adjusted correlation coefficient for 1967-68 or 1968-69 and  $r_2$  is that for 1967-69.

Table 2 contains a set of correlation coefficients based on those of

<sup>4</sup> For a derivation of eq. (5) see the Appendix.

Table 1 but subjectively adjusted to fit the above preconceptions. It also shows as  $Z_1$  and  $Z_2$ , respectively, the coefficients of  $X_{1968}$  and  $X_{1967}$  in equation (5). In addition, it compares the precision of an estimate based on  $X_{1968}$  alone with that of one using both  $X_{1968}$  and  $X_{1967}$ . The quantity  $V_1$  is the error variance  $1 - r_1^2$  (i.e., the portion of the total variance not "accounted for" by the correlation), which applies when credibility  $r_1$  is given to  $X_{1968}$ , and  $V_2$  is the error variance of the estimate given by equation (5) and is calculated from the following formula:<sup>5</sup>

$$V_2 = \frac{1 + 2r_1^2r_2 - 2r_1^2 - r_2^2}{1 - r_1^2}.$$

A comparison of  $V_1$  and  $V_2$  suggests that, under the particular circumstances, relatively little increase in precision is gained by using two years of claim experience instead of one.

#### IV. A BRIEF COMPARISON WITH PREVIOUS THEORIES OF CREDIBILITY

Previous theories of credibility<sup>6</sup> have generally shared the following characteristics:

1. The observed claim rate for a group is assumed to be distributed randomly about some other rate or quantity, sometimes called the "true" or "underlying" rate and sometimes even less explicitly defined.
2. The true rate for each group is taken to be constant for all years.
3. For appropriate distributions of the observed rate and of the true rate, one obtains the well-known formula for the credibility factor

$$Z = \frac{P}{P + K}, \quad (6)$$

where  $P$  is proportional to the exposure and  $K$  is a constant.

The approach taken in this paper does not share the characteristics listed above—in fact, this writer would reject all three.

The notion that the hypothetical true rate is constant seems to be inconsistent with some well-known facts, such as the fact of employee turnover; the existence of epidemics and other short-term but nonrandom

<sup>5</sup> Derived from Hoel, *op. cit.*, p. 115, eq. (8).

<sup>6</sup> See in particular, A. W. Whitney, "The Theory of Experience Rating," *PCAS*, IV (1918), 274; R. Keffer, "An Experience Rating Formula," *TASA*, XXX (1929), 130; A. L. Bailey, "Credibility Procedures," *PCAS*, XXXVII (1950), 7; A. L. Mayer-son, "A Bayesian View of Credibility," *PCAS*, LI (1964), 85; and R. D. Maguire, "An Empirical Approach to the Determination of Credibility Factors," *TSA*, XXI (1969), 1.



changes in morbidity and mortality; the relationship of morbidity and perhaps also of mortality to economic conditions; changes in local conditions of medical care, in prices and utilization; and observed long-term, nonrandom changes for the United States population as a whole and, for jumbo groups, changes of a magnitude too great to be explained as statistical fluctuation.

This leaves open the possibility of a true claim rate which changes in time. But a changing true rate which is not identical with the observed rate would be by definition unobservable, and no evidence could ever be adduced to prove its existence. Moreover, if the true rate is not constant but changes in unknown ways, there seems no point in trying to estimate its value; it is of no practical significance to us.

Regrettably, the approach of this paper leads to no simple formula such as equation (6) for the credibility factor. However, never have appropriate parameters been determined for equation (6) or for other formulas produced by previous theories, and the formulas have not been successfully applied in practice to group coverages.

Most of the previous theories were developed within the context of casualty insurance, while this one was developed for group insurance. There would, however, seem to be no essential impediment to the applicability of this theory to casualty coverages.

#### ACKNOWLEDGMENTS

The author wishes to acknowledge the helpful comments of Mr. John K. Kittredge and Mr. Erwin A. Rode, who read a preliminary draft of this paper, and the thorough review and many suggestions of Mr. William Katcher.

#### APPENDIX

##### DERIVATION OF EQUATION (5)

Let  $X_3$ ,  $X_2$ , and  $X_1$  denote loss ratios in three successive years. (Here the notation will differ from that used in the body of the paper, in that  $X_1$  is the most recent year.) Assume that the means and variances of all three variables are identical and that they are transformed so that they are measured from the mean.

Let  $r_{ij}$  be the correlation coefficient between  $X_i$  and  $X_j$ . Assume that  $r_{12} = r_{23}$ . Find  $a_2$  and  $a_3$  in the formula  $E(X_1) = a_2X_2 + a_3X_3$ . According to Hoel (p. 113),

$$a_i = -\frac{s_1R_{1i}}{s_iR_{11}},$$

238 CREDIBILITY OF GROUP INSURANCE CLAIM EXPERIENCE

where  $s_i$  is the standard deviation of  $X_i$  and  $R_{ij}$  is the cofactor of  $r_{ij}$  in the determinant

$$R = \begin{vmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{vmatrix}.$$

Because of the previous assumptions,  $a_i$  and  $R$  can be simplified:

$$a_i = -\frac{R_{1i}}{R_{11}}, \quad R = \begin{vmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{12} \\ r_{13} & r_{12} & 1 \end{vmatrix}.$$

Then

$$R_{11} = 1 - r_{12}^2, \quad R_{12} = -r_{12} + r_{12}r_{13}, \quad R_{13} = r_{12}^2 - r_{13},$$

$$E(X_1) = \frac{r_{12} - r_{12}r_{13}}{1 - r_{12}^2} X_2 + \frac{r_{13} - r_{12}^2}{1 - r_{12}^2} X_3.$$

This is tantamount to equation (5).

## DISCUSSION OF PRECEDING PAPER

ERNEST A. ARVANITIS:

This paper is a very welcome addition to the literature. I was particularly pleased to see such a paper, and I hope that it may serve as an impetus to bring together the practical actuary and the theoretical statistician in a unified attack on problems of estimation in group life and health insurance. The problems I am thinking of involve stop-loss premiums, credibility, and general risk theory.

The results that Mr. Margolin has reported for comprehensive major medical seem to support our own results. In studying the mean and standard deviation by size of group, we have discovered that for the group life coverage the standard deviation seems to continue to decrease as the size of group increases. The rate at which the standard deviation decreases with increase in size of group seems to be reasonably consistent with what might be expected if chance fluctuations were the major contributing factor. In the case of group health coverages, on the other hand, the decrease seems to be at a somewhat lower rate, and the standard deviation seems both to tend toward a plateau and to show substantial variation between years even for the largest groups. These characteristics are not consistent with the theory that chance fluctuations are the major factor involved. Other factors, such as economic conditions or the impossibility of accurately evaluating the risk beforehand, seem to be operating to a considerable degree.

Theoretical statistics should provide a very helpful tool if it is directed against the realities of the situation. In the past this theory has been limited in basically assuming that a random process was all that was involved. Given this basic premise, with all sorts of assumptions with regard to independence, considerable effort has been devoted to reducing the results to elegant mathematical terms which were of limited value in group life insurance and of almost no value in group health insurance. I would like very much to see practice and theory brought together, because neither by itself is of very much value. The statistician should devote more effort to explaining his assumptions in detail, together with the limitations involved, and should relegate his mathematical manipulations to appendixes. The actuary should devote more effort to quantifying his assumptions in some manner. Together, hopefully, they can then arrive at some agreement with respect to what constitutes an appropriate model. Perhaps we can then develop some procedures with which

to measure or take into account biases or other external factors. For example, is there some way of determining bias or poor estimation in the manual rate? Another approach is to ask whether the combination of manual rate, mean, standard deviation, and correlation are what one might expect from chance fluctuations. If not, can we perhaps bring to bear some tools such as the analysis of variance to measure or minimize the extraneous factors?

WALTER SHUR:

Mr. Margolin has presented a very interesting paper which is certain to create very vigorous discussion. Although his correlation approach is, in my judgment, a very natural and a very useful one, he makes a number of statements which will surely appear as red flags to any genuine Bayesian. It is my feeling that the difference between Mr. Margolin's approach and the typical Bayesian approach is more one of degree, albeit an important degree, than one of substance.

I believe that what Mr. Margolin is saying can be summed up briefly as follows:

Suppose that we are observing a very large block of cases exposed during 1969 and 1970 and that we consider these cases homogeneous for rate-making purposes. We have a particular case which had a 65 per cent loss ratio in 1970. How should we estimate the loss ratio on this case in 1971?

Mr. Margolin's method is based on the premise that one can do no better than to extract all the cases that had a 65 per cent loss ratio in 1969, determine what the average loss ratio for these cases was in 1970, say, for example, 78 per cent, and use this as the estimated loss ratio for the case in 1971 (provided, of course, that there is a sufficient volume of data). The relationship between the 65 per cent and the 78 per cent takes into account a myriad of factors, including pure statistical correlation which depends primarily on size, changes in age distributions, changes in product mix, changes in distributions of benefit amounts, vagaries of the manual rate structure, and the like. It is Mr. Margolin's contention that these various items are so numerous and complex that any attempt to build a realistic theoretical model is doomed to failure and will do more harm than good by imposing unrealistic conditions on the data.

The remainder of Mr. Margolin's method simply makes use of linear equations, as a matter of convenience, to predict next year's claim level. In the case of two variables this equation turns out to be the simple credibility formula  $ZX_{n+1} + (1 - Z)\bar{X}$ , where  $Z$  is equal to the usual correlation coefficient  $\rho$ .

The fact is that Mr. Margolin has himself created an a priori model, although it is a very simple one. This model, which is essential for obtaining his formulas, involves the following assumptions:

1. The average loss ratio for all cases in the homogeneous block is constant over the period of time required, for example, for 1969, 1970, and 1971 in the illustration given above.
2. The variance of the loss ratios for the block of cases is constant over the same period of time as in assumption 1.
3. The correlation coefficient between loss ratios in 1968 and 1969 is the same as that between loss ratios in 1969 and 1970 (in the three-variable case).
4. In obtaining credibility factors from the data, raw correlation coefficients are forced into a preconceived notion of uniform grading by size of case.

The most distinguished feature of Mr. Margolin's model is the absence of any assumption as to the form of any distribution function, and it is this lack that most sharply distinguishes his model from the typical Bayesian model.

The Bayesian approach makes predictions based on a blend of a priori knowledge and actual facts. It seems to me that the question Mr. Margolin has raised, and I think it is a good question, is whether our a priori knowledge as to the mathematical form of the loss-ratio distribution is good enough to be taken into account at all. He suggests that the complex practical situation is such that we are better off to obtain a sufficiently large volume of actual data and make our predictions almost entirely on the basis of factual relationships demonstrated in those data. Knowing something about the complexities of the group medical care business, I think that Mr. Margolin makes an excellent point. However, I still think that we can call Mr. Margolin a Bayesian, although, in deference to the simplicity of his a priori model, perhaps in his case we should spell it with a small "b."

The remainder of my discussion is primarily mathematical, except for a few closing comments concerning the quality of the mathematical derivation in the paper. The principal results are set forth in Sections I-V and are supported by the mathematical derivations which follow in Appendix Sections A-D.

#### *I. The Correlation Approach Is Not Inconsistent with the Bayesian Approach or with Mr. Maguire's Empirical Approach*

Under the correlation approach, the estimate of next year's claim level for a particular case is given by

$$x_2 = \rho x_1 + (1 - \rho)\mu, \quad (1)$$

where  $\rho$  is the correlation coefficient between the variables  $x_1$  and  $x_2$ ,  $x_1$  is this year's claim level, and  $\mu$  is the average claim level. (In practice, of course,  $\rho$  and  $\mu$  must be estimated from sample data.)

Under certain simple normality assumptions<sup>1</sup> concerning the prior distribution of "true" claim levels for all cases, and the distribution of actual claim levels on a particular case, it is shown mathematically in this discussion that  $\rho = 1/(1 + \sigma_A^2/\sigma_T^2)$ , and formula (1) becomes

$$x_2 = \frac{1}{1 + \sigma_A^2/\sigma_T^2} x_1 + \left(1 - \frac{1}{1 + \sigma_A^2/\sigma_T^2}\right) \mu, \quad (2)$$

where  $\sigma_A^2$  is the variance of the actual claim level and  $\sigma_T^2$  is the variance of the true claim level.

Thus, under these assumptions, the attempt to estimate  $\rho$  from the sample is simply an attempt to estimate  $1/(1 + \sigma_A^2/\sigma_T^2)$ . Similarly, I showed in my discussion of Mr. Maguire's paper that, under the same assumptions, his method also produced an estimate of  $1/(1 + \sigma_A^2/\sigma_T^2)$ .

In the case of more variables, that is, where next year's claim level  $x_{n+1}$  is to be estimated from previous years' claim levels,  $x_1, x_2, \dots, x_n$ , by means of multiple linear least-squares regression, the above assumptions lead mathematically to the formula

$$x_{n+1} = \frac{n}{n + \sigma_A^2/\sigma_T^2} \frac{x_1 + x_2 + \dots + x_n}{n} + \left(1 - \frac{n}{n + \sigma_A^2/\sigma_T^2}\right) \mu. \quad (3)$$

Formula (3) agrees precisely with formula (8) which appears in my discussion of Mr. Maguire's paper.

II. *Mr. Margolin's Correlation Approach Produces a Very Simple Formula, Regardless of the Number of Variables, if We Assume that  $\rho_{x_i x_j} = \rho$  for Each Pair of Variables*

In particular, we obtain the formula

$$x_{n+1} = \frac{n\rho}{1 + (n-1)\rho} \frac{x_1 + x_2 + \dots + x_n}{n} + \left[1 - \frac{n\rho}{1 + (n-1)\rho}\right] \mu, \quad (4)$$

where  $\rho$  and  $\mu$  must be estimated from the observed values of  $x_1, x_2, \dots, x_n$  (see Sec. III below for a suggested method of making these estimates).

It should be noted that formula (4) does not depend on the normality assumptions referred to in Section I above; the only assumption required is that  $\rho_{x_i x_j} = \rho$  for each pair of variables. Also, it might be noted that formula (5) in Mr. Margolin's paper reduces to formula (4) above if we set  $r_1 = r_2$  in the former and  $n = 2$  in the latter.

<sup>1</sup> The same assumptions that I made in my discussion of R. Maguire, "An Empirical Approach to the Determination of Credibility Factors," *TSA*, XXI (1969), 1. They are spelled out later in the mathematical derivations in the Appendix to this discussion.

Formula (4) is interesting in that it expresses the credibility factor in terms of the correlation coefficient, as  $n\rho/[1 + (n-1)\rho]$ , which varies between 0 and 1 as  $\rho$  varies between 0 and 1.

III. *The Following Formulas Are Suggested for Estimating  $\mu$  and  $\rho$  from the Observed Values; They Were Obtained by the Method of Maximum Likelihood, under the Normality Assumptions Referred to in Section I Above*

Suppose that we are observing  $N$  cases and are looking at the experience of these cases in each of  $n$  observation periods. Let  $x_{il}$  be the claim level (loss ratio) observed in the  $i$ th period on the  $l$ th case. Then, under the normality assumptions spelled out in Section I above, the maximum likelihood estimates of  $\mu$  and  $\rho$  are obtained from the following natural formulas:

$$\hat{\mu} = \frac{1}{nN} \sum_{i=1}^n \sum_{l=1}^N x_{il}; \quad (5)$$

$$\hat{\sigma}^2 = \frac{1}{nN} \sum_{i=1}^n \sum_{l=1}^N (x_{il} - \mu)^2; \quad (6)$$

$$\hat{\sigma}_{x_i x_j} = \frac{1}{n(n-1)N} \sum_{i=1}^n \sum_{j=1}^n \sum_{l=1}^N (x_{il} - \mu)(x_{jl} - \mu), \quad i \neq j; \quad (7)$$

$$\hat{\rho} = \frac{\hat{\sigma}_{x_i x_j}}{\hat{\sigma}^2}. \quad (8)$$

IV. *Even if One Does Not Accept the Normality Assumptions Referred to in Section I Above, the Estimate ( $r_{xy}$ ) of  $\rho$  in Mr. Margolin's Method Should Be Made on the Basis of the Formulas in Section III Above*

Mr. Margolin's derivation of the formula  $x_2 = \rho x_1 + (1 - \rho)\mu$  assumes that  $\mu_{x_2} = \mu_{x_1} = \mu$  and  $\sigma_{x_2} = \sigma_{x_1}$ . The need for this assumption is evident if we begin with the well-known formula for the least-squares regression line of  $x_2$  on  $x_1$ , namely,

$$x_2 - \mu_{x_2} = \rho \frac{\sigma_{x_2}}{\sigma_{x_1}} (x_1 - \mu_{x_1}). \quad (9)$$

Setting  $\mu_{x_2} = \mu_{x_1} = \mu$ ,  $\sigma_{x_1} = \sigma_{x_2}$  in formula (9) leads directly to

$$x_2 = \rho x_1 + (1 - \rho)\mu.$$

Therefore, when the sample is used to estimate  $\rho$ , the formula for the estimate should reflect the assumptions made about the parent distribution. This is done in formulas (5)–(8), which produce unique estimates

of  $\mu_1\sigma^2$  and  $\sigma_{x_r x_s}$ , each of which is based on all the available data in all the observation periods, and then combine these estimates to produce an estimated  $\rho$ . In contrast, the standard method used in the paper employs an  $\bar{x}_1$  and an  $\bar{x}_2$  (which will undoubtedly differ) and a  $\sigma_{x_1}^2$  and  $\sigma_{x_2}^2$  (which will undoubtedly differ) in the calculation of  $\sigma_{x_1 x_2}$  and  $\rho$ . This unnecessary additional variability simply makes the estimate of  $\rho$  less efficient (that is, the distribution of the estimate about the true value has a greater variance).

*V. An Alternative Correlation Assumption Also Provides Some Strong Theoretical Underpinning for the Simple Formula  $x_n = \rho x_{n-1} + (1 - \rho)\mu$*

The formula  $x_n = \rho x_{n-1} + (1 - \rho)\mu$  used for estimating next year's claim level from this year's actual level and the average level has a remarkable interpretation if we postulate that the claim levels on a particular case over a period of  $n$  years are correlated by means of the relationship  $\rho_{i,j} = \rho^{|i-j|}$ . That is, the correlation between any two loss ratios is  $\rho$  raised to a power equal to the number of years separating them.

Under that assumption, the application of multiple least-squares regression to estimate  $x_n$  from  $x_1, x_2, \dots, x_{n-1}$  produces the surprising result that  $x_n = \rho x_{n-1} + (1 - \rho)\mu$ ; that is, the coefficients of  $x_1, x_2, \dots, x_{n-2}$  turn out to be zero. Hence the simple  $x_n = \rho x_{n-1} + (1 - \rho)\mu$  can (under certain theoretical assumptions) be interpreted as a case of multiple linear regression. Of course, the estimate of  $\mu$  would involve all the previous years' loss ratios, not just  $x_{n-1}$ .

If we also assume that the loss ratios  $x_1, x_2, \dots, x_n$  are governed by a multivariate normal distribution, the support for  $x_n = \rho x_{n-1} + (1 - \rho)\mu$  is even greater, since the least-squares hyperplane would be, in fact, the curve of regression. In this instance, the conditional distribution of  $x_n$  given  $x_1, x_2, \dots, x_{n-1}$  depends only on  $x_{n-1}$ .

As should be apparent from the first part of this discussion, I do not mean to imply from the relationship between the credibility and correlation approaches that the correlation approach is therefore not new. Quite the contrary—I believe that the correlation approach (which might more properly be called the least-squares approach to cover the case of more than two variables) introduced by Mr. Margolin is a natural one in its own right and makes a great deal of sense as a basic rationale for making statistical estimates of future loss ratios. The best of all possible functions for estimating  $y$  from  $x$ , in a least-squares sense, is the curve of regression—that is, the average value of  $y$  for a particular value of  $x$ . From an insurance point of view, the average would seem to be the



desired index. Linear regression simply produces (for convenience) a linear approximation to the curve of regression.

Finally, it is important to keep in mind a clear separation between theory and practice. From a business point of view, Mr. Margolin's paper and this discussion are theoretical, notwithstanding the taking of samples and estimating of parameters. The financial management of the group business involves global estimates of aggregate renewal premium required for blocks of business. A system for estimating future loss ratios on individual cases should be looked at primarily as a model which helps to distribute the aggregate renewal burden equitably among all cases.

#### *Appendix: Mathematical Derivations*

Because of their importance in what follows, I repeat the assumptions made in my discussion of Mr. Maguire's paper, although in somewhat briefer form.

Suppose that we are dealing with a very large block of cases for which the following assumptions are reasonable:

1. The true claim levels  $\tau$  for the various cases in this large block are normally distributed about the known average claim level for the block,  $\mu$ , with variance  $\sigma_T^2$ .
2. The actual claim level which emerges on a particular case in the block, during any single observation period, is normally distributed about the true claim level  $\tau$  for that case, with variance  $\sigma_A^2$ .

#### *A. Joint Distribution Function for $x_1, x_2, \dots, x_n$*

Let  $x_1, x_2, \dots, x_n$  be random variables which represent the claim levels on a single case in  $n$  observation periods. Then, under the normality assumptions given above,

$$\begin{aligned}
 f(x_1, x_2, \dots, x_n) &= \int_{-\infty}^{\infty} \frac{1}{\sigma_T \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{\tau - \mu}{\sigma_T}\right)^2\right] \frac{1}{\sigma_A \sqrt{2\pi}} \\
 &\quad \times \exp\left[-\frac{1}{2} \left(\frac{x_1 - \tau}{\sigma_A}\right)^2\right] \dots \frac{1}{\sigma_A \sqrt{2\pi}} \\
 &\quad \times \exp\left[-\frac{1}{2} \left(\frac{x_n - \tau}{\sigma_A}\right)^2\right] d\tau.
 \end{aligned} \tag{10}$$

When the exponential functions are combined, they will contain the expression

$$S = \left(\frac{\tau - \mu}{\sigma_T}\right)^2 + \frac{1}{\sigma_A^2} \sum_{i=1}^n (x_i - \tau)^2.$$

If we (i) rewrite  $(x_i - \tau)$  as  $(x_i - \mu) - (\tau - \mu)$ , (ii) expand the summation, (iii) complete the square on terms involving  $(\tau - \mu)^2$  and  $(\tau - \mu)$ , and (iv) define  $\rho$  and  $\sigma$  so that  $\rho\sigma^2 = \sigma_T^2$  and  $\sigma^2(1 - \rho) = \sigma_A^2$ , and substitute for  $\sigma_T$  and  $\sigma_A$  nearly everywhere, then  $S$  can be algebraically transformed into

$$S = \left( \frac{\tau - K}{\sigma_A \sqrt{\rho/[1 + (n - 1)\rho]}} \right)^2 + \frac{1}{\sigma^2(1 - \rho)[1 + (n - 1)\rho]} \times \left\{ [1 + (n - 2)\rho] \sum_{i=1}^n (x_i - \mu)^2 - \rho \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n (x_i - \mu)(x_j - \mu) \right\},$$

where  $K$  is a constant of no importance.

Noting that the product of coefficients in expression (10) can be written as

$$\begin{aligned} \frac{1}{\sigma_T \sigma_A^n (2\pi)^{(n+1)/2}} &= \frac{1}{\sigma_A \sqrt{\rho/[1 + (n - 1)\rho]} \sqrt{2\pi}} \frac{\sqrt{\rho/[1 + (n - 1)\rho]}}{\sigma_T \sigma_A^{n-1} (2\pi)^{n/2}} \\ &= \frac{1}{\sigma_A \sqrt{\rho/[1 + (n - 1)\rho]} \sqrt{2\pi}} \frac{1}{(2\pi)^{n/2} \sigma^n \sqrt{[1 + (n - 1)\rho](1 - \rho)^n}} \end{aligned}$$

and that

$$\begin{aligned} &\frac{1}{\sigma_A \sqrt{\rho/[1 + (n - 1)\rho]} \sqrt{2\pi}} \\ &\times \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2} \left\{ \frac{\tau - \mu}{\sigma_A \sqrt{\rho/[1 + (n - 1)\rho]}} \right\}^2 \right] d\tau = 1, \end{aligned}$$

we have finally from equation (10) that

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sigma^n \sqrt{[1 + (n - 1)\rho](1 - \rho)^{n-1}}} e^{-Q_n/2}, \tag{11}$$

where

$$\begin{aligned} Q_n &= \frac{1}{\sigma^2[1 + (n - 1)\rho](1 - \rho)} \\ &\times \left\{ [1 + (n - 2)\rho] \sum_{i=1}^n (x_i - \mu)^2 - \rho \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n (x_i - \mu)(x_j - \mu) \right\}. \end{aligned}$$

As a consequence of the definitions in step (iv) above, we have

$$\rho = \frac{1}{1 + \sigma_A^2/\sigma_T^2}. \tag{12}$$

It will now be shown that equation (11) gives the standard multivariate normal distribution, where  $\rho$  is, in fact, the correlation coefficient between each pair of variables  $x_r$  and  $x_s$ . This, together with equation (12), will be proof of equation (2).

The multivariate normal distribution, where  $\sigma$  is the standard deviation of each variable and  $\rho$  is the correlation coefficient between each pair of variables, is given by<sup>2</sup>

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sigma^n \sqrt{\Delta}} e^{-Q_n/2}, \quad (13)$$

where

$$Q_n = \sum_{i=1}^n \sum_{j=1}^n \frac{\Delta_{ij}}{\sigma^2 \Delta} (x_i - \mu)(x_j - \mu),$$

$\Delta$  is the  $n \times n$  determinant of the correlation coefficients  $\rho_{ij}$ , and  $\Delta_{ij}$  is the cofactor of  $\rho_{ij}$  in  $\Delta$ . In the present instance,

$$\Delta = \begin{vmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \rho & \rho & 1 & \dots & \rho \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \rho & \rho & \rho & \dots & 1 \end{vmatrix}.$$

It can be shown by induction that

$$\Delta = [1 + (n-1)\rho](1-\rho)^{n-1}; \quad (14)$$

$$\Delta_{ii} = [1 + (n-2)\rho](1-\rho)^{n-2}; \quad (15)$$

$$\Delta_{ij} = -\rho(1-\rho)^{n-2}, \quad i \neq j. \quad (16)$$

Substituting these values in equation (13), we obtain precisely the same expression for  $f(x_1, x_2, \dots, x_n)$  as in equation (11), showing that equation (11) gives, in fact, a multivariate normal density, and  $\rho$  is, in fact, the correlation coefficient between each pair of variables.

### B. Multiple Linear Least-Squares Regression

In order to estimate  $x_{n+1}$  from  $x_1, x_2, \dots, x_n$  by means of linear least-squares regression, we must find the coefficients  $\beta_i$  by the method of least squares, in

$$(x_{n+1} - \mu) = \beta_1(x_1 - \mu) + \beta_2(x_2 - \mu) + \dots + \beta_n(x_n - \mu). \quad (17)$$

<sup>2</sup> H. Cramér, *Mathematical Method of Statistics* (Princeton, N.J.: Princeton University Press, 1951), sec. 24.2.1.

The value of  $\beta_i$  is given by<sup>3</sup>

$$\beta_i = - \frac{\Delta_{(n+1)i}}{\Delta_{(n+1)(n+1)}}, \quad (18)$$

where  $\Delta_{(n+1)i}$  and  $\Delta_{ii}$  are as defined in Section A above but  $\Delta$  is an  $(n+1) \times (n+1)$  determinant. Using equations (15) and (16), we have

$$\beta_i = - \frac{-\rho(1-\rho)^{n-1}}{[1+(n-1)\rho](1-\rho)^{n-1}} = \frac{\rho}{1+(n-1)\rho}. \quad (19)$$

Substituting equation (19) into equation (17) and rearranging terms proves equation (4). Note that this derivation did not require the normality assumptions set forth above. Under those normality assumptions, we have already shown in Section A that  $\rho = 1/(1 + \sigma_A^2/\sigma_T^2)$ . Making that substitution in equation (4) proves equation (3).

### C. Maximum Likelihood Estimators

Under the normality assumptions given above,  $f(x_1, x_2, \dots, x_n)$  is given by equation (11). Assume now that we are observing  $N$  cases in each of  $n$  observation periods, and let  $x_{il}$  be the claim level on the  $l$ th case in the  $i$ th observation period. The likelihood function  $L$  is given by

$$L = \left[ \frac{1}{(2\pi)^{nN/2} \sigma^{nN} \{ [1+(n-1)\rho](1-\rho)^{n-1} \}^{N/2}} \right] \exp \left( - \frac{1}{2} \sum_{i=1}^N Q_n \right),$$

where  $Q_n$  is given by equation (11) with  $x_i$  and  $x_j$  replaced by  $x_{il}$  and  $x_{jl}$ , respectively, summed over  $l$ . Formulas for the estimators are determined by differentiating  $\ln L$  in turn with respect to  $\mu$ ,  $\sigma$ , and  $\rho$ , setting each of the three results equal to zero, and solving simultaneously.

To simplify the writing in what follows, the following symbols will be used:

$$\sum_{l,i} = \sum_{l=1}^N \sum_{i=1}^n (x_{il} - \mu)^2;$$

$$\sum_{l,i,j} = \sum_{l=1}^N \sum_{i=1}^n \sum_{j=1}^n (x_{il} - \mu)(x_{jl} - \mu), \quad i \neq j;$$

$$\begin{aligned} \ln L = & - \ln C - nN \ln \sigma - \frac{N}{2} \ln [1 + (n-1)\rho] \\ & - \frac{N(n-1)}{2} \ln (1-\rho) - \frac{[1+(n-2)\rho] \sum_{l,i} - \rho \sum_{l,i,j}}{2\sigma^2 [1+(n-1)\rho](1-\rho)}. \end{aligned}$$

<sup>3</sup> *Ibid.*, sec. 23.2.4.

Differentiating  $\ln L$  with respect to  $\mu$ , and setting the result equal to zero, we find that  $\rho$  and  $\sigma$  vanish, leading directly and easily to equation (5).

Differentiating  $\ln L$  with respect to  $\sigma$ , we easily obtain the following:

$$\hat{\sigma}^2 = \frac{[1 + (n - 2)\rho] \sum_{l,i} - \rho \sum_{l,i,j}}{nN[1 + (n - 1)\rho](1 - \rho)}. \quad (20)$$

Differentiating  $\ln L$  with respect to  $\rho$ , and setting the result equal to zero, we obtain the following equation:

$$\begin{aligned} \frac{Nn(n - 1)\rho}{2(1 - \rho)[1 + (n - 1)\rho]} - \frac{\rho\{[1 + (n - 1)\rho](n - 2) + n\}}{2\sigma^2[1 + (n - 1)\rho]^2(1 - \rho)^2} \sum_{l,i} \\ + \frac{1 + \rho^2(n - 1)}{2\sigma^2[1 + (n - 1)\rho]^2(1 - \rho)^2} \sum_{l,i,j} = 0. \end{aligned} \quad (21)$$

Multiplying through by the denominator of the last two terms, and replacing  $\sigma^2 Nn(1 - \rho)[1 + (n - 1)\rho]$  by its value obtained from equation (20), equation (21) becomes

$$\begin{aligned} (n - 1)\rho \left\{ [1 + (n - 2)\rho] \sum_{l,i} - \rho \sum_{l,i,j} \right\} \\ - \rho \{ [1 + (n - 1)\rho](n - 2) + n \} \sum_{l,i} + [1 + \rho^2(n - 1)] \sum_{l,i,j} = 0. \end{aligned} \quad (22)$$

Equation (22) reduces readily to

$$\hat{\rho} = \frac{\sum_{l,i,j}}{(n - 1) \sum_{l,i}}, \quad (23)$$

which proves equation (8).

From equation (23) we have

$$\sum_{l,i,j} = (n - 1)\rho \sum_{l,i}.$$

Substituting this in equation (20) and simplifying proves equation (6). Equation (7) was defined only for convenience, to show the naturalness of the formula for  $\rho$  and as an aid in remembering that formula.

#### D. Alternative Correlation Assumption, $\rho_{ij} = \rho^{|i-j|}$

Suppose that  $x_1, x_2, \dots, x_n$  are random variables which represent the claim levels on a single case in  $n$  observation periods, with  $x_n$  for the most recent period,  $x_{n-1}$  for the next prior period, and so on. We assume

further that  $\sigma_{x_i} = \sigma$  and  $\mu_{x_i} = \mu$ , but at this point we make no assumptions as to the form of the distribution function for  $x_1, x_2, \dots, x_n$ .

The determinant of the correlation coefficients is given by

$$\Delta = \begin{vmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \rho^2 & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \rho & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \rho^{n-4} & \dots & 1 \end{vmatrix}.$$

It can be shown (although all of these results are not needed in what follows) that

$$\begin{aligned} \Delta &= (1 - \rho^2)^{n-1}; \\ \Delta_{11} &= \Delta_{nn} = (1 - \rho^2)^{n-2}; \\ \Delta_{ii} &= (1 + \rho^2)(1 - \rho^2)^{n-2}, \quad i = 2, 3, \dots, n - 1; \\ \Delta_{ij} &= 0, \quad |i - j| \neq 1, \quad i \neq j; \\ &= -\rho(1 - \rho^2)^{n-2}, \quad |i - j| = 1. \end{aligned} \tag{24}$$

Following the method described in Section B, but recognizing that we are dealing with  $n$  variables rather than  $n + 1$ , and that the time order of variables is different, we have

$$x_n - \mu = \beta_{n-1}(x_{n-1} - \mu) + \beta_{n-2}(x_{n-2} - \mu) + \dots + \beta_1(x_1 - \mu).$$

Since  $\beta_i = -\Delta_{ni}/\Delta_{nn}$ , we have, from equation (24),

$$\beta_{n-1} = \rho, \quad \beta_{n-2} = \beta_{n-3} = \dots = \beta_1 = 0.$$

Hence  $x_n - \mu = \rho(x_{n-1} - \mu)$ , and  $x_n = \rho x_{n-1} + (1 - \rho)\mu$ .

It is indeed a remarkable result to "throw away" all but the most recent year's claim experience. This theoretical result was confirmed to me by Professor Hickman, who noted that others have been troubled by it and no satisfactory explanation has yet been offered.

Finally, I feel obliged to comment on the quality of the mathematical derivation in the paper. In general, some of the notation and definitions are so inconsistent with accepted practice as to greatly confuse anyone knowledgeable in the subject; there are several incorrect formulas, and the entire derivation does very little more than develop the standard equation  $y - \bar{y} = r_{xy}(\sigma_y/\sigma_x)(x - \bar{x})$  for a least-squares line.

To be specific, the paper defines the expression  $E(X_n)$  to be the estimated value of the random variable  $X_n$ . The symbol  $E(X_n)$  is standard, of course, for the mean of the variable  $X_n$ . The integral in equation (1) is immediately recognizable as  $\sigma_{X_n}^2$ —but it is not, in consequence of the paper's unusual definition. Further, the paper defines  $\text{cov}(a, b)$  as the expected value of the product  $ab$ , whereas traditionally the term covariance is used for the expected value of  $[(a - \bar{a})(b - \bar{b})]$ .

Footnote 3 defines  $\text{cov}(a, b)$  as

$$\int_{-\infty}^{\infty} abf(a)da .$$

This is incorrect even for the definition intended in the paper and should be a double integral, namely,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} abf(a)f(b)dadb .$$

Equation (4) in the paper is also incorrect, since  $\text{cov}(X_n, X_{n-1})$  is not really the covariance, as explained in the preceding paragraph. The reason that the right conclusion is reached is simply that the mean of  $X_n$  is equal to the mean of  $X_{n-1}$  (say,  $\mu$ ), and therefore the equation  $X_n = X_{n-1} + e_n$  can be replaced by  $(X_n - \mu) = (X_{n-1} - \mu) + e_n$ . This essential assumption concerning the means is not even stated in the derivation.

The entire derivation can be replaced by the following: The standard equation for a least-squares line, for the regression of  $X_n$  on  $X_{n-1}$ , is  $X_n - \mu_{X_n} = \rho(\sigma_{X_n}/\sigma_{X_{n-1}})(X_{n-1} - \mu_{X_{n-1}})$ , where  $\rho$  is the correlation coefficient between  $X_n$  and  $X_{n-1}$ . Assuming that  $\mu_{X_n} = \mu_{X_{n-1}} = \mu$ , and  $\sigma_{X_n} = \sigma_{X_{n-1}}$ , and substituting in the above equation for the least-squares line, we obtain  $X_n = \rho X_{n-1} + (1 - \rho)\mu$ , the result the paper is seeking to develop.

WILLIAM J. SCHREINER:

Mr. Margolin is to be congratulated for his fresh and elegant approach to the search for a satisfactory predictor of a group insurance plan's future claim rate. In addition, in a very few words he very effectively delivers telling blows to the classical theory of credibility. The utility of any mathematical model is measured by the degree to which it approximates the world it seeks to represent, and he makes what I believe to be correct observations on the differences between classical credibility theory and the world of group insurance.

While my admiration for this paper is considerable, I find that it is

subject to the same criticism it levels at classical credibility theory, since it gives us no reason to believe that the methodology it presents will develop a more satisfactory model for predicting future claim experience than that which it seeks to replace. Mr. Margolin's approach assumes that (1) past experience is useful for predicting future experience on a particular group, (2) average expected experience on all groups of the same type is useful for predicting future experience on a particular group, and (3) a linear combination of past experience and average expected experience is a satisfactory predictor of future experience. In his paper, however, Mr. Margolin suggests that neither past experience nor average expected experience, by themselves, is a satisfactory predictor of future experience. Furthermore, he suggests that, aside from computational convenience, there is nothing to recommend a linear combination of these items with respect to the prediction of future results. If one agrees with these observations (and I do), one is, I feel, compelled to ask whether this method gains any ground toward solving the question at hand.

Lest one be discouraged by these observations, however, I would like to suggest that the ability to predict a particular group's future experience is not a prerequisite for the insurer who seeks a successful financial result. While this statement may seem paradoxical, I am sure it will be reassuring to those of us who have ever faced the task of predicting a satisfactory premium level for a given case in the face of the myriad factors, unknown and perhaps unknowable, that will be operating both internally and externally to determine the case's future experience.

The key factors in developing a proper premium level are (1) the distribution of the actual claims about the predicted claims, irrespective of the method that is used to obtain the predicted claims, and (2) the insurer's dividend formula with respect to the given class of policyholders.

Without going into a complete development of the rationale involved, it will be helpful to note that the insurer's basic objective is to obtain sufficient premium so that, when interest earnings are added and incurred claims, expenses, and dividends are subtracted, a satisfactory positive gain results.

With this in mind, let us consider a two-case portfolio of a particular insurer who has fairly good luck in predicting what the portfolio's aggregate claims will be but has the misfortune to estimate 10 per cent too high on one case and a corresponding 10 per cent too low on the other. Further, assume that the insurer would like to have a \$2,000 gain at the end of the first year and that it will incur \$10,000 of expenses in connection with each case. It is also assumed that no interest will be earned.



The figures in Table 1 suggest that, if no dividends were to be paid, a proper premium would be \$61,000 for each case, as the desired gain would result as shown in Table 2.

Next, let us consider a related and perhaps more practical situation in which, under the same experience conditions, the insurer utilizes a dividend formula which refunds premium to policyholders with favorable experience. In particular, let us assume that the insurer returns 80 per cent of the "unexpected" claim saving to the policyholder from whom

TABLE 1

	Case No. 1	Case No. 2	Total
Desired gain.....	\$ 1,000	\$ 1,000	\$ 2,000
Incurred expenses.....	10,000	10,000	20,000
Expected claims.....	50,000	50,000	100,000
Total.....	\$61,000	\$61,000	\$122,000

TABLE 2

	Case No. 1	Case No. 2	Total
Earned premium.....	\$61,000	\$61,000	\$122,000
Incurred claims.....	45,000	55,000	100,000
Incurred expenses.....	10,000	10,000	20,000
Gain (loss) before dividend.....	\$ 6,000	(\$ 4,000)	\$ 2,000
Dividend.....	0	0	0
Gain (loss) after dividend.....	\$ 6,000	(\$ 4,000)	\$ 2,000

it originates. Assuming that it is not known which case will produce the better-than-expected experience, we see that, to achieve the same \$2,000 net gain, the insurer must increase his premium charge by \$2,000 for each case, as shown in Table 3.

Since each of the foregoing examples was based on an identical ability to predict an individual case's future claim experience, it becomes evident that the key element in determining the premium level required to achieve the insurer's gain objective was not the predictive ability of the insurer; rather, it was the dividend formula selected by the insurer. Other dividend methods would have resulted in different conclusions with respect to the proper premium level. Similarly, had the deviation in actual claim

experience from the expected been different, still other premium conclusions would have been reached in order that the insurer's gain objective might be achieved.

In addition to suggesting that knowledge of the interrelation between the distribution of actual claim results about the expected and the insurer's dividend formula is the primary ingredient for a profitable result, I believe that the examples indicate that a highly developed ability to predict future claim experience for a given policyholder is neither required, nor sufficient by itself, for the insurer to achieve financial success. This, in turn, suggests that it really does not matter, from a practical

TABLE 3

	Case No. 1	Case No. 2	Total
Earned premium.....	\$63,000	\$63,000	\$126,000
Incurred claims.....	45,000	55,000	100,000
Incurred expenses.....	10,000	10,000	20,000
Gain (loss) before dividend.....	\$ 8,000	(\$ 2,000)	\$ 6,000
Dividend.....	4,000	0	4,000
Gain (loss) after dividend.....	\$ 4,000	(\$ 2,000)	\$ 2,000

standpoint, whether we ever find the perfect credibility factor. While this may be troubling to the theorist, since the search has gone on for over forty years without apparent success, I find it extraordinarily reassuring.

## JAMES C. HICKMAN:

One of the universal problems shared by all branches of actuarial science is that of modifying the price-benefit structure of an insurance system as actual experience is revealed. Unfortunately, each branch has tended to adopt its own special nomenclature when designing a rational procedure for blending the information generated by a particular risk or line of business with the information obtained from past or ancillary experience or from more inclusive classifications of risks. Consequently, the fact that this adjustment process is a theme that unifies much of actuarial science is often obscured.

Credibility is an idea that seems to have had its genesis among North American casualty actuaries. In the most recent of a long series of papers on credibility that have appeared in the *Proceedings of the Casualty*

*Actuarial Society*, C. C. Hewitt [5] defines credibility as "a linear estimate of the true (inherent) expectation derived as a result of a compromise between hypothesis and observation." Hewitt, in common with most previous authors on this subject, illustrates credibility formulas in which the credibility factor is a function of the number of trials or exposure units. This has been considered essential, for traditionally the credibility factor has been viewed as a weight for use in computing a modified estimate of the expected value of the loss index under consideration. In this computation the revised estimate becomes the weighted average of the recent experimental value of the claim index and the previous estimate. It has seemed self-evident, in constructing a model to facilitate the discussion of this problem, that the weights (credibility factors) should depend on some measure of the size of the claim experience.

The author of this paper is also concerned with extracting useful information from the record of recent claim experience. By assumption, however, he limits his analysis to classes characterized by having approximately the same risk size: his model does not contain parameters that measure the size of a risk. The objective is the analysis of the time series of recent claim indexes for the purpose of short-term prediction.

In the author's basic model it is assumed that the random vector of claim indexes  $(X_1, X_2, \dots, X_{n+1})$  has a multinormal distribution. Then many perplexing problems are swept away by assuming that the  $n + 1$  random variables have a common mean and a common variance. Normal distribution theory then tells us that the conditional expected value of  $X_{n+1}$  given the other variables is the linear function  $\Sigma a_i(x_i - \mu)$ , where summation is from 1 to  $n$ ,  $\mu$  is the common mean, and the constants  $a_i = -\Lambda_{n+1,i}/\Lambda_{n+1,n+1}$ ,  $i = 1, 2, \dots, n$ ;  $\Lambda_{ij}$  is the  $i, j$  cofactor in the determinant of variances and covariances [4, p. 315].

The author might have developed his main result without assuming the multinormal distribution if his linear estimate had been viewed as a least-squares estimate of the expected value of  $X_{n+1}$  given the previous values of the claim index. If he had stressed this approach, his results, except for the suppression of parameters relating to risk size, would belong to the general family of credibility formulas developed by using least-squares linear approximations to conditional expected values developed by Bühlmann [2, chap. 4].

Although one may quibble about whether the time series analysis of claim data for the purpose of price-benefit structure modification is strictly a subset of credibility, the author's central point that actuaries should learn more about practical time series analysis is well taken. An

overview, from the executive level, of currently fashionable forecasting tools is provided by Chambers, Mullick, and Smith [3]. The pullout summary sheet in their article is its most valuable component. On a more operational level, the book by Box and Jenkins [1] is a rich source of ideas and examples.

To supplement the paper with an additional small taste of the analysis of time series, let us briefly examine a model that yields forecasts somewhat like those provided by equation (5) in the paper. We let  $X_t$  denote a claim index random variable for year  $t$ , and  $e_t$  ( $t$  a positive integer) denote a member of a set of mutually independent random variables with identical normal distributions, each with mean zero and a common positive variance. We adopt the model

$$X_t = X_{t-1} + (e_t - \theta e_{t-1}), \quad |\theta| < 1,$$

$$\Delta X_{t-1} = (1 - \theta E^{-1})e_t,$$

where  $\Delta$  is the finite-difference operator,  $E$  is the displacement or shift operator, and  $\theta$  is a parameter that dampens the impact of the past random shocks to the process. The autocorrelation function for this process, the correlation coefficient between  $X_{t+k}$  and  $X_t$  for  $k$  an integer, is denoted by  $\rho_k$ . It can be shown that

$$\begin{aligned} \rho_k &= 1, & k &= 0, \\ &= -\theta/(1 + \theta), & k &= 1, \\ &= 0, & k &\geq 2. \end{aligned}$$

In fact, one of the ways of identifying this model would be to compute sample autocorrelation functions and compare them with  $\rho_k$ .

If our objective is prediction, we might consider the conditional expected value of  $X_{t+1}$  given the values of the previous loss indexes. Our prediction would be given by

$$(1 - \theta)[X_t + \theta X_{t-1} + \theta^2 X_{t-2} + \dots].$$

Of course, an estimate of  $\theta$  would have to be used in an application. This is the familiar exponential smoothing formula which has been suggested frequently as a way to solve sales forecasting problems.

In this discussion we have attempted to develop and illustrate the author's idea that time series analysis may be useful in analyzing actuarial data. However, there are some built-in conceptual problems when these methods are used to analyze claim indexes that are expressed in terms of deviations from expected results.

By what standard do we judge a classification system and a price-benefit structure? The best answer appears to be that success has been achieved when deviations from expected results are independently distributed. If dependencies exist in the sequence of deviations, these dependencies might be exploited by one of the parties to the insurance contract. In an efficient market, or a completely equitable market, characterized by complete information and open market determination of prices, the force of competition would tend to remove such dependencies. This concept of market efficiency, as related to independent deviations, is at the heart of the random walk hypothesis about speculative prices. The correlation coefficients that the author computed might also be used to test the effectiveness of the classification and price-benefit structure adjustment mechanism. The fact that they are positive for the years he studied, rather than distributed in a narrow band around zero, comes as no surprise when one considers the powerful economic forces that have existed in recent years. These forces have caused premium adjustments to lag behind changes in the economics of the health system.

Even if each of the correlation coefficients computed by the author were near zero, indicating a removal of dependencies within the price-benefit index under study, the problem of estimating the expected claim rate, unfortunately labeled  $\bar{X}$  in the paper, remains. In performing this estimation, a growing body of methods for incorporating prior and ancillary information, as well as directly relevant recent claim data, is being developed. In the dynamic economy in which we operate, the rewards for wringing the last drop of insight from the body of current information for the purpose of forecasting, and perhaps controlling, future results are enormous.

#### REFERENCES

1. BOX, G. E. P., and JENKINS, G. M. *Time Series Forecasting and Control*. San Francisco, Calif.: Holden-Day, 1970.
2. BÜHLMANN, HANS. *Mathematical Methods in Risk Theory*. New York: Springer-Verlag, 1970.
3. CHAMBERS, J. C., MULLICK, S. K., AND SMITH, D. D. "How to Choose the Right Forecasting Technique," *Harvard Business Review*, Vol. XLIX (July-August, 1971).
4. CRAMÉR, H. *Mathematical Methods of Statistics*. Princeton, N.J.: Princeton University Press, 1946.
5. HEWITT, C. C. "Credibility for Severity," *Proceedings of the Casualty Actuarial Society*, Vol. LVII (1970).

HERBERT L. FEAY:<sup>1</sup>

The information available to this participant in the discussion indicates that the premium charges for group health insurance are not determined by "generally accepted actuarial principles." The paper provides support for this conclusion. A scientific study using acceptable standards requires that the total experience for any collection of risks be separated into and be analyzed for each of the significant and measurable underwriting classifications represented in the group.

The paper first discusses manual premiums which are presumably used to determine the total aggregate premiums for the first policy year for a group contract. The paper states that manual premiums are "one's a priori estimate of future claims plus loadings for expenses and contingencies" and are based on "only very imperfect knowledge of what has actually happened." The comment is made, further, that the development of manual premium rates includes "much personal judgment" and that the group actuary or group underwriter will "rewrite the manual and invent subjective new rules to fit what he regards as unprecedented." We read further that manual rating schemes are "in essence compilations of personalist or subjective probabilities with only limited bases in fact."

After the first year, the average aggregate manual premium rate secured for the first year is modified by the actuary, using past experience in accordance with "the degree of credibility he attaches to the actual experience." The procedure appears to be based on aggregate loss ratios determined as the ratio of total incurred claim benefit payments to the total manual premiums for one year for a group policy. The expected loss ratio is formulated as an a priori estimate of future claim experience. The formulation involves personal opinion and subjective probabilities. There is no indication that there is a detailed determination of the expected future total claim costs using actuarially sound classifications of risks and probabilities supported by actual experience. The actual loss ratio is based on the total estimated incurred claim experience. Each loss total (expected or actual) is divided by the total manual premiums to secure the loss ratio (expected or actual). The period for determining these loss ratios is usually one year.

The accuracy (or credibility) of these aggregate ratios is indefinite and indeterminate. Obviously these ratios will vary from year to year for the same policy and also from policy to policy in the same year. These variations result from measurable differences in the exposure of the risks for acceptable actuarial risk classifications as well as from changes in

<sup>1</sup>The views expressed in this discussion are those of the participant and do not necessarily represent the views of the agency in which he is employed.

the basic probabilities and averages subject to credibility determination. The ratios for these aggregate totals also are distorted because the manual premiums include an allowance for expenses and the claim totals cover benefit payments without any addition for expenses.

Obviously, personal opinion and subjective probabilities have a large influence in the determination of premiums for group health insurance. Agency pressure, competition, and demands for premium volume must have considerable effect in the selection of premium rates when these rates are not supported by actuarially determined probabilities and averages secured from adequate and accurate classifications of statistics for exposures and claims.

Insurance laws specify that policyholders of a company are to receive equitable treatment. One wonders whether equity for the cost of insurance is maintained among the group health insurance contracts of a company when subjective probabilities have a substantial role in the determination of premiums collected. For example, small groups usually do not have the same pressure ability as large groups. However, there is little or no published information that will provide support for a factual statement on this question of equity among the groups. Not only do premiums lack statistical support, but the retrospective rating formulas used for group health insurance are not available to the public.

The apparent purpose of the secrecy is to limit information available to competitors, but this does not prevent the changing of groups from one carrier to another. A substantial number of group insurance contracts in force on any specified date will be transferred to another company at some time thereafter. Such changes are discouraged by law for individual insurance but seem to be an accepted practice for group insurance.

In view of the lack of adequate statistics to support premium charges, it is not surprising that a considerable number of companies incur losses from group health insurance operations.

An article in a recent insurance publication reports on the premium difficulties for a group health contract issued to a unit of government. The premium rates for the group were so understated that the policyholder has been asked to agree to a substantial premium increase in the middle of the contract year. A member of the judiciary, in discussing the case, explained that the carrier for the group was "in effect explaining it is customary for insurance agents to miss the first time around when they've got somebody on the hook."

The credibility of the basic aggregate averages that are used for projections is statistically indeterminate. Because of this, the use of correlation ratios for two successive years also gives results for which the credi-

bility is indeterminate. This is indicated by some of the comments in the paper. The correlation coefficients of Table 2 are "based on those of Table 1 but subjectively adjusted to fit the above preconceptions." This appears to be equivalent to adjusting the statistical results to agree with preconceived opinions.

One of the reasons given for the changes in the ratios of Table 1 is that incurred claim costs used for calculating aggregate loss ratios are inaccurate because of unreliable estimates for unpaid claims. Much of this inaccuracy is due to failure to use proper actuarial techniques for incurred but unpaid claims and failure to adjust cost estimates to actual results as the experience develops. The writer of this discussion has an actuarial note [2] on estimating incurred claims that provides a procedure for determining the present value of outstanding unpaid claims. The procedure uses as a base the amount of claims incurred and paid in the same year and ratios for incurred but unpaid claims based on past experience. The writer of this discussion has used a modification of this procedure for group health insurance. The principal change is separation of the unpaid claims by month of payment rather than by year of payment. The past experience for group health insurance shows a rapid reduction in the monthly total for delayed claims as the period of delayed payment increases, and indicates that all incurred claims for a year are paid within about 30 months after the end of the year. With the estimate of outstanding unpaid claims made at the end of 3 months after the close of a year, the estimate is for about 27 months for the last complete year and is for about 15 months for the next-to-the-last complete year. This writer will be pleased to give details of his method upon request to anyone interested.

Credibility is a statistical measure, and, if dependable credibility ratios are to be secured, accurate classifications of basic reliable statistics must be used. The scientific procedures and principles for statistically measuring credibility are the same as those for determining stop-loss reinsurance premiums, risk charges, and surplus fund limits. The first requirement is that the premium rates be established at accurate levels in accordance with generally accepted actuarial standards, taking into consideration acceptable risk classifications (such as inflation, age, sex, geographical location, and income) for which sufficient statistical information is available for determination of reliable premium rate differentials. These requirements are basic, so that the difference in actual experience from the estimated experience will be caused primarily by random fluctuations.

These scientific procedures and principles are covered by the subject called "risk theory." The writer of this discussion has taken part in



preparing two papers on this subject [1, 3], and the remarks that follow are based on those papers.

The two most important statistical problems in the application of risk theory are the determination of the expected mean claim cost per unit of exposure for a specified classification of risks and the measurement of the variation in that cost. These problems are essentially sampling problems, with each period of insurance (usually a year) for each classification being a sample of the experience for the very large generalized collection of similar risks. Average claim costs per unit of exposure will vary during one year among like groups and will vary for the same group from year to year because of random fluctuations. The causes of these random fluctuations are large in number, are substantially independent, and are unpredictable, and the variations resulting from each cause are a relatively small part of the total variation.

There are causes of variations in average claim costs per unit of exposure for a group of risks that do not meet the requirements for random fluctuations. Among those causes are the following: (1) age and sex, (2) occupation, (3) geographical location, (4) long-term changes in mortality and morbidity rates, (5) monetary inflation, and (6) concentration of risks in a limited area (catastrophe hazard).

Actuarial procedures should be used that will either eliminate or greatly reduce the effect of such causes of variation, so that the effects of each cause on claim cost variations become those of random fluctuations. The first requirement is, of course, satisfactory actuarial investigations to determine the exact effect of the larger measurable and predictable causes in average claim costs per unit of exposure. The results, combined with an accurate census of a group, can be used to determine an accurate estimate of the claim costs for the group.

Although the actuarially determined premium rates based on acceptable underwriting classifications (age, sex, income, and so on) may be the same, the average premium rates can vary considerably from group policy to group policy and from year to year for the same policy because of variations in the underwriting classifications for the persons insured. An example is an increase in the percentage of persons at ages over 65 years.

The individual health insurance study published in the 1969 *Reports* number of the *Transactions of the Society of Actuaries* illustrates the subdivision of a large group of risks into reliable underwriting classifications. Classifications of this kind are needed for group insurance if actuarial determinations are to be made of credibility, stop-loss reinsurance premiums, risk charges, and surplus requirements.

These comments apply to any broad grouping of individual risks,

regardless of whether they are issued individual policies or are issued certificates for insurance under a group policy. The actuary studying the experience must give consideration to the proper classifications of individual risks with similar underwriting characteristics included in the total "heterogeneous lot."

One example of the disregard of this actuarial principle would be studies of average health costs by state. These studies are used to produce aggregate average costs per unit of exposure without regard to the other underwriting factors. For example, there is nothing to show the effect on the higher aggregate average costs sometimes listed for California and Florida of possibly more mature populations in those states as compared with, say, Kansas and Iowa.

The rate-makers for private group health insurance need to apply the motto of the Society and "substitute facts for appearances and demonstrations for impressions." The business needs scientifically accurate premiums, risk charges, and surplus limitations if private group health insurance is to be advocated as a partial answer to providing health care at a reasonable cost to the people of the nation.

There is reference in the paper to the first edition of *Introduction to Mathematical Statistics* by P. G. Hoel, copyrighted in 1947. The edition is difficult to locate and has been succeeded by a second edition, copyrighted in 1954, and by a third edition, copyrighted in 1962.

#### REFERENCES

1. FEAY, HERBERT L. "Introduction to Nonproportional Reinsurance," *Transactions of the Society of Actuaries*, XV, 94.
2. FEAY, HERBERT L. "Discussion of Estimating Incurred Claims," *ASTIN Bulletin*, V (Part II), 274.
3. FEAY, HERBERT L., AND KABAK, IRWIN W. "Frequency Formulas for Determining Stop-Loss Reinsurance Premiums," *Journal of Risk and Insurance*, XXXV, 371.

#### (AUTHOR'S REVIEW OF DISCUSSION)

MYRON H. MARGOLIN:

I wish to thank the five discussants for their stimulating comments. Each writes from a unique point of view and upon different facets of the total credibility problem. They also raise a number of new questions.

Mr. Arvanitis' remarks on the relationship between theory and reality and his skepticism toward elaborate but untested mathematical models express articulately the empirical approach required to solve problems of the real world. His discussion is also a brief but excellent example of the "scientific method": he reports his observations of the data, draws

whatever generalizations they imply, and only then begins to speculate on what kind of mathematical model (if any) might be appropriate.

I am grateful to Mr. Shur for a number of reasons. First, in Sections III and IV of his discussion, he points out that the maximum likelihood principle produces better estimates of the means, variances, and correlations required by my approach than I had used in the paper. Second, he has replaced my derivation of equation (4) by a simpler, more rigorous one. Third, he has drawn some interesting mathematical comparisons between my approach and previous approaches under certain hypothetical assumptions. Fourth, he has concisely summarized the a priori assumptions inherent in my "model."

Mr. Schreiner's remarks on mathematical models and on the classical theory of credibility clearly place him in the empiricist camp together with Mr. Arvanitis and me. However, while I agree with him that there is an interrelationship between an insurer's renewal rating formula and his dividend formula, I think he overstates the connection when he suggests that the latter formula fully compensates for any deficiencies in the former. The worse the experience rating formula, the larger will be the deviations—both plus and minus—between the actual claims and those anticipated in the rates. The larger these deviations, the higher must be the risk charges levied in the dividend formula against the cases with positive results to make up for the losses on the others. This is both inequitable and noncompetitive.

Mr. Shur makes somewhat the same point when he terms my paper "theoretical" and goes on to say that the real objective is to distribute the aggregate renewal burden equitably. It seems to me that the proper way to achieve equity is to minimize the deviations between actual and expected claims, which in turn minimizes the need for some groups to subsidize others. As in classical statistics, it is a practical necessity to minimize the sum of the squares of the deviations rather than the absolute values of the deviations themselves. I should add that my approach has been put to practical use within my own company.

In their discussions Mr. Shur and Professor Hickman refer frequently to "models." This is not strange. As actuaries, we deal with complex social-financial phenomena in mathematical terms, often by constructing a model into which we build what we consider to be the essential characteristics of these phenomena. A mortality table is a familiar example of a model. Most efforts to solve the problem of credibility have also used models. Accordingly, the balance of this discussion will be devoted to various models proposed for group insurance credibility, including the model implicit in Mr. Feay's remarks.

I should first disclose a fairly strong personal bias against elaborate mathematical models, unless they are buttressed with strong empirical support. There exists always the danger that a model will completely omit one or more significant or essential features of whatever it is supposed to represent. For example, a model airplane does not disclose that a real one flies people from place to place. The conventional mortality table is useful for premium calculations, but it omits completely the fact that mortality rates change in the course of time, and it is useless (by itself) for examining long-term mortality trends.

Persons who do not share my bias may believe it proper to advance models on purely theoretical or a priori grounds or on the basis of general reasoning, without any empirical support. To persuade others to accept their models is normally much more difficult without data. It would not take a large quantity of data to convince most scientists that the height of adult American males is (hypothetically) approximately normally distributed with mean and standard deviation of 5 feet 10 inches and 4 inches, respectively, and increasing by 2 inches per generation; but to convince them that they should accept this simple mathematical model solely on a priori grounds—that is, to try to prove on the basis of genetics, geography, nutrition, and so on, that height *should* be normally distributed and *should* increase—is obviously a practical impossibility.

In constructing a model, it is best to include as few unsubstantiated preconceptions as possible. The more one tells the model, the more it echoes the voice of the model-builder. In one of the discussions it was suggested that my model might better have included some parameters reflecting size of risk. If this means that I should have made some assumptions as to how loss ratios behave or how credibility factors should vary by size, I cannot agree. If my approach is valid, then the numbers in Table 2 tell us how credibility factors actually do vary by size, given the particular manual premium rating system and actual loss ratios studied. These numbers speak for themselves rather than for the model-builder.

Mr. Shur has concisely stated the four assumptions in my "model." The first, that the average loss ratio is the same in all years, is equivalent mathematically to a simple transformation of variable; in practice, it means that you want to estimate the aggregate trend factor correctly. The second assumption, that the variance of the loss ratios is the same in all years, will not in practice be realized precisely. To illustrate what might cause the variance to change, suppose that the claim costs in New York City were increasing more rapidly than the national average, and in Houston less rapidly. If the manual premiums do not respond to these

changes, then the variance should increase. New York groups should tend to show progressively higher costs in relation to the average loss ratio, and Houston groups progressively lower. This sort of distortion cannot be corrected, in my opinion, by any conceivable credibility system but only by the manual premium rating system. The third assumption is that the correlation between successive years' loss ratios will be the same as in the past.

These assumptions strike me as fairly plausible, but no one has to take them on faith. They can be empirically confirmed or disproved. The data in Table 1 of my paper lend at least mild support to them.

The fourth assumption, that the correlation coefficients should progress smoothly by size, is the same kind of preconception that underlies the graduation of any set of statistical data, such as a mortality table.

Professor Hickman has explained privately his statement that my model assumes a multinormal distribution for the loss ratios. He agrees that no distribution need be assumed for equation (5), provided that I am willing to accept  $E(X_{1969})$  as a least-squares approximation to the conditional expected value of  $X_{1969}$  and not as the conditional expected value itself. This I willingly accept. My approach is nothing more than the application of the least-squares principle with no assumptions made for the distributions of the loss ratios. In fact, it is doubtful whether there exists a stable mathematical distribution of these loss ratios and hence whether there exists a unique conditional expected value for a future loss ratio.

In considering the other models discussed below, we should keep in mind that (at least to my knowledge) no empirical evidence has ever been adduced in support of any of them. The discussion must therefore be rather theoretical, inquiring whether their assumptions are clearly stated, are not inconsistent with what we know of the real world, and are free from logical self-contradiction.

Mr. Feay asserts that if group actuaries used what he terms "generally accepted actuarial principles," we could produce the perfect manual premium rating system for group health insurance—that is, one in which the deviations between actual and expected claims were largely or entirely statistical fluctuations. He apparently believes that what I shall call the "urn-wager" model is an appropriate representation of group insurance claim experience. According to this model, each group case corresponds to an urn filled with numbered balls. The composition of each urn is different from that of the others. One ball is drawn for each person in the group, and the sum of the numbers drawn corresponds to the number of claims in a year. The amount of each claim may correspond

to further selections from another set of urns. The premium charged is in effect a wager on the total value of the balls drawn. (To be sure, Mr. Feay did not explicitly refer to such a model, but it is conceptually and mathematically identical with the notion that, after adjusting for age, sex, and so on, the differences between actual and expected claims are random.) By properly analyzing our experience, breaking it down into the right classifications, group actuaries can allegedly estimate fairly accurately the composition of the urn (i.e., the true claim rate) based on seven or so parameters describing the group (age, sex, occupation, and so on).<sup>1</sup>

In practice, group actuaries do precisely what Mr. Feay asks of them. We do analyze our experience by age, sex, and similar variables, but such analyses cannot possibly lead to the objective, foolproof rating system he envisions. Consider only two of the variables—geography and industry. The decision of how finely to subdivide the country into geographical regions and what these regions shall be is clearly subjective. The Standard Industrial Classification Listing contains 78 major industries subdivided into hundreds of industry groups, which are in turn subdivided into close to 2,000 industries. Suppose that a particular group actuary subjectively designates 100 regions and 200 industry groups, or 20,000 cells in all. Then most of the individual cells will contain only one or two small groups, hardly deserving of much credibility. Deciding how to combine them further into meaningful risk classes entails only more subjectivity.

Suppose that a particular cell of ten cases contains nine small ones and one larger than the other nine combined. Shall the actuary include; exclude, or dampen the effect of the large one? What if the large one is in a central city but many of the others are in its suburbs?

After establishing cells with a credible number of groups in each cell, the actuary finds that no two benefit plans in a cell are quite alike. Furthermore, certain types of benefit plans tend to be favored by certain types of industries or in certain geographical areas, making it difficult to distinguish the effects of plan from geography or industry.

In order to establish separate geographical and industry factors, the actuary may look at the experience of all the cases in one geographical area. To get the right geographical factors he must first know the proper industry factors, but to get the right industry factors he must know the

<sup>1</sup> Actually, only four of the parameters he lists characterize individual groups. These are age, sex, occupation, and geography. Two others deal with area or nationwide trends. The seventh, concentration of risks, is not related to a group's expected claim rate but perhaps is related to its degree of fluctuation.

right geographical factors. This type of problem can be attacked by using multivariate analysis techniques, but these do not solve the problem nearly so precisely as envisioned by Mr. Feay. The results, which might be in the form of a multiple regression equation, do not generally give precisely the "right" prediction for any cell. They typically reduce the error variance by no more than 25 per cent.

In addition to these practical difficulties, the urn-wager model is wrong in principle. An urn is static. Its contents do not change. The attributes of its contents which determine the sample results are enumerable and measurable (that is, one is able to count how many balls are marked with each number). It does not interact with its environment. The same characteristics do not apply to a group of people, to their rates of disability, or to their expenses for medical care.

Income is an additional factor often used in manual rating but omitted by Mr. Feay. This is suggestive of the important role played by economic factors in health care costs.

Considered on a national basis, predictions of health care costs are much like predictions of the gross national product. The health care system is a major component of our economic system, accounting for nearly 10 per cent of the gross national product. Like economic activity in general, health care costs are subject to inflation, to the effects of supply and demand, and to rapid technological, sociological, and political change.

Much of consumer spending, whether in general or for medical care in particular, is for the satisfaction of basic physical needs. How the consumer satisfies these needs—when and to whom he goes to satisfy them and how much he is willing to spend—is usually his choice, influenced by innumerable economic, social, and psychological variables. Other spending is discretionary. Much illness is psychosomatic, and some surgery is elective. There is much evidence that hospital utilization and the incidence of even "nonelective" surgery can be greatly altered by altering the economic incentives to patients and providers. Particularly in the areas of disability, dental care, psychiatric treatment, cosmetic surgery, and maternity rates, there is abundant evidence that age, sex, geography, and occupation do not nearly tell the whole story.

If the problem could be solved by finding the frequency of illness for various risk classes and the distribution of their claim expenses per illness, then estimating the gross national product must be essentially a matter of estimating the frequency with which various types of people will work and the frequency distribution of their wages. (The right analogy may be "wage-earner," not "urn-wager.")

If it is difficult to estimate next year's health care expenses for the nation at large, it is even harder to estimate the health care expenses for a particular group. Predicting the costs of one tiny part of the system re-

quires an understanding of the system as a whole but in addition a detailed knowledge of what distinguishes this tiny part from the rest. Knowledge of age, sex, geography, and other factors is helpful only to a degree. Consider the following examples:

During one recent year, the number of days of disability per thousand home office employees was twice as high in one large life insurance company as in another. The companies are located only a few miles apart, their age-sex-occupation mixes are similar, and the numbers of employees involved are so great as to warrant essentially full credibility.

The claim costs under the comprehensive major medical coverage for a medium-sized company in rural Illinois shot up dramatically from one year to the next. It turned out that an interstate highway had just been opened, facilitating travel to a city where hospital rates were about twice those of the local hospital.

The experience of numerous groups in the aircraft and aerospace industries has shown that, during periods of layoffs or when layoffs seem imminent, claims under most health insurance coverages skyrocket. Employees who are laid off or about to be laid off attempt to obtain maximum benefit from their medical and dental coverages before they run out, and employees who have worked in spite of potentially disabling conditions suddenly become "objectively" disabled.

When the major hospitals of a large city simultaneously increase their rates by a percentage substantially higher than that assumed in one's premium rates, the resulting effect on the major medical experience of all groups in and near the city is hardly a "random" fluctuation.

The extent to which "sampling" enters into the problem can be gauged by the following examples:

At any given time, a certain percentage of males aged 35 are in the hospital. If sufficiently large random samples are drawn from the nation at large at one-minute intervals, the percentage in the hospital will be about the same at successive intervals, but tomorrow may be a Saturday or Sunday, and all economic activity, including hospital care, tends to lessen over the weekend. Three months later the percentages will differ because of "seasonal factors" (another term drawn from economics). A year from now the percentages may be still different. Hospital occupancy rates have been lower in 1971 than in 1970, and hospital administrators attribute this drop to general economic conditions.

Suppose that one compared the claim experience among 35-year-old male clerical employees of Company A with that of Company B, the two companies being located in the same geographical area and in the same industry. No a priori reason can be given why they should be the same. By definition of the term "random sample," the employees of Company A and those of Company B are not random samples drawn from the same population. Companies do not



hire on a random basis, nor do employees randomly choose where to work. Working conditions and personnel practices may differ substantially. The companies may be several miles apart and tend to attract employees from different neighborhoods, hence from different socioeconomic groups.

The weight of the evidence indicates rather strongly, I believe, that group health insurance costs cannot be accurately predicted and that the problem is too complex to be reduced to a mathematical model in which deviations between actual and expected are largely random. The burden of empirical proof lies with those who assert that an accurate rating system is possible.

From this discussion, one might infer that group health care expenses are not insurable risks. Professor Hickman observes that an imperfect manual rating system leaves the door open to antiselection. He conjectures that in an efficient market the force of competition should bring us close to the perfect manual system envisioned by Mr. Feay, one in which actual claims deviate only randomly and independently from those anticipated by the manual rates. In practice, insurers recognize the practical if not the theoretical impossibility of achieving a perfect manual rating system and try to minimize antiselection in a different way. To determine a suitable payable rate, insurance companies bidding on a case of moderate or large size try to obtain the claim experience with the present carrier. This reduces the buyer's advantage vis-à-vis the seller. Obviously, the risk to the insurer is greater when previous experience is not available on large groups.

Moreover, transfer of nonrandom risks is an integral part of our economic system. When a manufacturer bids on a contract at a fixed price, he is taking the risk that his costs may exceed his estimate. The financial relationship is essentially the same as a nonparticipating group contract, except that an insurer is estimating the buyer's claim costs instead of his own production costs. In some situations, where the risk seems too great or too difficult to measure, either the manufacturer or the insurer may seek to operate on a "cost-plus" basis.

Professor Hickman may be correct in saying that my approach resembles economic time series analyses, and perhaps actuaries should learn more about such analyses. However, this was not the "central point" of the paper, which was to propose a new approach to credibility. In dwelling upon the analogies with economics in this discussion, my purpose is to question whether any manual premium rating system or any credibility model can accurately capture all the essential features of the real world. Perhaps one can, but we should ask for empirical verification before accepting it, as we would of an economic model.

Professor Hickman illustrates the possible connection between time series analyses and credibility with a model in which the change in claim level from year to year depends on "random shocks." This example may have been an unfortunate choice. The correlation coefficient between  $X_{t+k}$  and  $X_t$  is zero for  $k > 1$ . A priori, it seems illogical that 1971 experience could correlate with 1970 experience and 1970 with 1969 without a correlation also, to some extent, with 1969. This preconception is confirmed by a self-contradiction in the model. Because the unit of time in which claim experience is measured is theoretically arbitrary, by picking shorter and shorter units (one year, one month, one week, one day), one is led to the conclusion that there is never any correlation between current and past experience.

It seems more reasonable to suppose that a model might faithfully represent group life experience. Unlike sickness, death is objective. Generally the individual has little choice in when he shall die and how much the claim will be.

Nevertheless, social, economic and psychological factors cannot be completely discounted. Suicide is obviously psychological in origin, and persons in some socioeconomic groups are much more likely than others to die of homicide. The incidence of heart attacks and the incidence of accidents are each influenced by personal "life style." There are significant seasonal and long-term changes in population mortality rates. The result of a recent study "clearly indicates that economic downturns are associated with increased mortality from heart disease and that, conversely, heart disease mortality decreases during economic upturns."<sup>2</sup>

In evaluating applications for individual life insurance, underwriters take account of numerous factors, many of them partly subjective, in addition to age, sex, and occupation. Whether these other factors somehow cancel out among a group of employees of the same employer is not obvious.

The accompanying tabulation shows the group life experience on one of

Policy Year Ending October 1	Average Lives Insured	Expected No. of Deaths	Actual No. of Claims	Claim Rate per 1,000	U.S. Unem- ployment Rate (Cal. Year)
1965 . . . . .	19,591	64.7	76	3.88	4.5%
1966 . . . . .	21,649	71.4	80	3.70	3.8
1967 . . . . .	24,620	81.2	78	3.17	3.8
1968 . . . . .	27,594	91.1	66	2.39	3.6
1969 . . . . .	30,577	100.9	82	2.68	3.6
1970 . . . . .	31,576	104.2	131	4.15	4.9
1971 . . . . .	30,798	101.6	113	3.67	5.9 (est.)

<sup>2</sup> M. H. Brenner, "Economic Changes and Heart Disease Mortality," *American Journal of Public Health*, LXI, No. 3 (March, 1971), 609.

our largest cases for the last seven years. Because reliable and complete age data are not available for most of the years, expected deaths were calculated on the basis of a constant annual rate of 3.30 per thousand. This rate almost precisely equalizes the deviations between actual and expected for 1968 and 1970. We have estimated that the probability of two random fluctuations of such magnitude and of opposite direction in a three-year period is less than 2 in 10,000. If it were also taken into account that, based on the growth in lives, the average age was (apparently) declining through 1970, this a priori probability would be even less. The progression of the experience also closely parallels national unemployment rates.

I think that the Bayesian models were proposed with group life or casualty experience in mind, not group health. There are three distinct types of Bayesian models on which I would like to comment, but they all share the following characteristics:

1. The actual loss ratio in year  $t$ ,  $X_t$ , is assumed to be conditionally distributed around an unobservable parameter  $\theta_t$ .
2. There is in general a different  $\theta_t$  for each group. The  $\theta_t$  for all groups in year  $t$  follows another assumed distribution about the mean  $\mu_t$ .
3. On the basis of the first two assumptions and of an observed value for  $X_t$ , one applies Bayes theorems to estimate  $\theta_t$  for that group.

A question applicable to all Bayesian models is: what happens when  $X_t$  and  $\theta_t$  are expressed in terms of different manual premium rating systems? If they were expressed in terms of pure claim dollars, then surely the  $X_t$  and  $\theta_t$  could be distorted from year to year and even within a year by large nonrandom changes in the number of lives or the age distribution. If they are expressed as loss ratios, then the appropriate distribution for  $X_t$  and  $\theta_t$  must be peculiar to the particular system used. No one model can have general validity.

The model cited by Mr. Shur is an example of that type of Bayesian model in which  $\theta_t$  is considered to be the group's "true" loss ratio and is held to be constant for all  $t$ . I suspect that, theoretically, the  $\theta_t$  can be constant in only one manual premium system and that in this system the value of  $\theta_t$  must be the same for all groups. Pure theory aside, this type of model is equivalent to the "urn-wager" model and hence is not even approximately valid for health coverages. Perhaps it would work for group life, but is there an empirical way to find out?

Bühlmann<sup>3</sup> has proposed a second type of model in which the  $\theta_t$  are "risk parameters"—apparently like true claim rates which are allowed to vary from year to year. He gives one example (p. 164) in which the

<sup>3</sup> H. Bühlmann, "Experience Rating and Credibility," *ASTIN Bulletin*, V (Part II) (1969), 157.

$\theta_t$  may increase by secular trend, but this example can be converted to a constant true loss ratio model by incorporating trend into the manual rates. In general, it seems to me that if the relationship between successive  $\theta_t$  is known, the model can always be converted into one of the first type; if the relationship is stochastic (changes are random but follow a known distribution), it is not possible to distinguish between these random changes in the true loss ratio and the random fluctuations of  $X_t$  around the true loss ratio; and if the relationship is unknown, the model is quite useless.

In the third type,  $\theta_t$  is some sort of subjective hypothesis, and the frequency distribution of  $\theta_t$  represents one's uncertainty about  $\theta_t$ . However, the meaning of  $\theta_t$  has never been adequately clarified, in my opinion. For example, suppose that  $\theta_t$  is subjectively given a normal distribution with a mean of 0.7 and a standard deviation of 0.1. What is it that is supposed to have a 95 per cent chance of lying between 0.5 and 0.9? Certainly it is not the actual loss ratio, since  $X_t$  has its own distribution around  $\theta_t$ . Furthermore, if  $\theta_t$  is subjective, then so must be the conditional distribution of  $X_t$  given  $\theta_t$ . Thus the specifics of the distributions are peculiar not only to one manual premium rating system but also to one person.

The only way to make logical sense of the model may be to assume that  $\theta_t$  is one's subjective estimate of a true objective loss ratio, as in the first two models. However, subjectivist Bayesians do not believe in objective probabilities or in objective true claim rates, so that the question of what their hypothesis refers to remains unclear.

Perhaps my own assumptions can be refuted, but it seems to me that they are about the weakest possible assumptions consistent with the notion that there is some regularity in how the real world behaves from year to year, without which an actuarial science and an insurance industry are not possible.