If all competing carriers were to use the same experience rating plan year after year, if the plan's predictive accuracy were lower than it need be, and if all underwriters were equally ignorant of how low that accuracy was; then nothing adverse would happen that an off-balance factor couldn't cope with. The three requisites might be difficult to maintain indefinitely, but while they were maintained the plan could be looked upon as one in which the policyholder pays some portion of last year's losses next year and thus is rewarded for being a good risk and punished for being a poor one. Only enough predictive accuracy is required to keep the third requisite in effect.

At the other extreme, there is a market model in which each carrier uses a plan different from that of every other carrier. More importantly, each policyholder is a perfect price buyer who considers each year's insurance as a separate transaction and annually shops the entire market for the lowest renewal quotation which he buys. Under such conditions, I believe all carriers having experience rating plans with less than the best predictive accuracy would be in financial difficulty. I'm not certain what would happen to the carrier whose plan had the best predictive accuracy. It might do satisfactorily or it might only be in less financial difficulty than the companies with less gifted actuaries. But in either event the overpowering demand for predictive accuracy would call for statistical models capable of using more and different kinds of information than the model we have been talking about.

Of course, such a concept of the market doesn't correspond with reality either. Other things being not too unequal, most buyers prefer to maintain a continuing relationship with the same carrier or producer. Service effectiveness and service satisfaction typically improve with time. Even price buyers tend to have more confidence in longer period comparisons than shorter period ones. A company that rates its business in such a way as to make its better customers feel at home should expect them to respond by staying there.

Is it possible to imagine a statistical model that has a closer correspondence to the pricing problem? In such a model predictive accuracy probably would not be controlling, but certainly conspicuous deficiencies of it would limit the inertia of the business. That inertia-like characteristic would be recognized quantitatively, together with the factors that contribute to it. The model should permit the buyer to dissociate himself from insurance pooling to the extent that he is willing to absorb his own losses, either directly or in rating. If a simpler definition of rateable losses brings about an easier meeting of the minds, the model should balance that gain against the loss of predictive efficiency. If the buyer wishes both to minimize pooling and to contain fluctuations, the model should permit him to extend himself in time.

It is easier to point to the elements that an existing statistical model does not contain than it is to design a better one. I am not at all certain that a better one can be designed or that one containing the elements I have mentioned would be a foundation for constructive mathematical inquiry. I do feel that Mr. Bailey could do it if anyone could.

DISCUSSION BY LEWIS H. ROBERTS

This paper is distinguished by two virtues which are unfortunately not often found in combination: on the one hand, incisive theoretical analysis, and on the other, thorough practicality. The first deserves mention because the author has developed the mathematical basis for a crucial problem which is not only one of our most difficult, but one which for many years has provided a broad field of uncertainty upon which the exponents of executive intuition have jousted in darkness with the upholders of underwriting judgment. The second virtue was demonstrated to this reviewer by the important applications in which the formulas presented in this paper have proved to be exceedingly valuable in competitive rate making.

The author gives as the first criterion for experience rating:

"Each dollar of loss, or absence thereof, should contribute to the risk's adjusted rate an amount equivalent to the amount of information it provides regarding the future losses of the same risk for the same amount of exposure."

The "amount of information" provided by a statistic is defined in statistical theory by the equation

$$(1) \qquad I = 1/\sigma^2$$

where σ is the standard deviation of the statistic.

This leads immediately to Gauss' theorem on observation weights, according to which observations with varying degrees of reliability, or precision, are averaged with weights equal to I to yield minimum variance, hence maximum information, for the average. Reliance upon this theorem is implicit in the author's basic approach, and provides the essential element of mathematical soundness which has unfortunately been the missing ingredient of more than one other treatment of this subject.

The author's second criterion, that the risk's premium should not fluctuate widely from year to year, appears to be self-evident. Generally, however, only if the experience of a risk is allowed to contribute *more* to the adjusted rate than the amount of information it contains—that is, if too much credibility is given to the experience—would wide fluctuations occur. We might regard this second criterion, therefore, as providing a symptom that the first has been violated.

His third criterion is that a dollar of actual loss should not add more than a dollar to the adjusted losses because otherwise the insured might find it to his advantage to pay his own losses. The author's reason might not be clear to everyone. There is no objection to the insured's paying his own losses under deductible and excess insurance. Why should there be any objection here? The answer seems to be that it is not so much the insured's *paying* his own losses that bothers us—it is his *not reporting* them. This third criterion, like the second, is essentially symptomatic of violation of the first. The problem here arises particularly when amounts of loss are ignored and only the number of claims is considered. It is not surprising that anomalies should result from such oversimplification. Rigorous adherence to the first criterion should preclude premium debits that exceed the amount of actual loss and expense, as the author's formulas demonstrate.

The fourth criterion, that an experience rating should not be too expensive to administer, is incontestable. It may be, however, that we actuaries allow ourselves to be too defensive about administrative costs, and tend to underestimate the profit value of an efficient pricing system for insurance.

One of the author's particularly trenchant remarks is that ". . . the Actuarial Theory of Indeterminacy . . . would state that when we get sufficiently refined statistics in sufficient volume to be able to determine the correct values for an experience rating plan, we won't use the information that way because we can then determine a far better class plan instead. It is when the data is limited and hence the rates less accurate that the need for experience rating is greater." Very true. Some time has been known to elapse, however, before the results of actuarial research found their way into a class plan, and then only after the defects of the plan had been profitably exploited by independent underwriters.

Without, of course, attempting to repeat the author's derivations in detail here, let us examine the salient features of his thesis. By way of introduction to the author's formulas, a few comments on their theoretical background may be appropriate.

We shall use the symbol $\underline{E}$ to denote expected values. The word "expected" will be used here only in the statistical sense of theoretical average.

In the sense of conserving the most information, the most efficient formula upon which to base the credibility of a risk's experience is given by:

(2) $$Z = \sigma_E^2 / (\sigma_A^2 + \sigma_E^2)$$

where A and E are the only available estimators of the inherent hazard, H, of a risk, and $\underline{E}(A - H) = \underline{E}(E - H) = 0$. If A is the risk's actual losses, and the inherent hazard remains unchanged since the experience period, we define H as equal to $\underline{E}A$, hence $\underline{E}(A - H) = 0$. Where a risk with inherent hazard H is chosen at random from a class of risks with average inherent hazard E, it follows for any such choice that $\underline{E}H = E$, hence $\underline{E}(E - H) = 0$. (It is important to note that where experience rating is optional it is incorrect to regard H as a random choice, hence we are not justified in assuming $\underline{E}(E - H) = 0$ in discussing optional plans.)

The variance of actual losses, σ_A^2, is by definition $\underline{E}(A - H)^2$. We define σ_E^2 as $\underline{E}(E - H)^2$, which is necessarily equal to $\underline{E}(H - E)^2$ or σ_H^2, the latter being the variance of the class and a measure of its heterogeneity. Note that σ_E^2 is not used here to denote the variance of E, which has no variance, but σ_E^2 is the mean square error associated with the use of E to estimate each H. It is therefore identically equal to the variance of H.

A theoretical weakness of most, if not all, experience rating plans in use today is that credibility cannot be measured in accordance with Eq. (2) above because σ_A^2 is reflected only roughly and σ_E^2 is ignored altogether. The importance of taking σ_E^2 into account is pointed up by the efficiency of so-called "merit" rating plans in which substantial discounts and surcharges are soundly developed for risks whose experience would have no credibility whatever under traditional experience rating formulas.

The author implicitly applies Eq. (2) with important results when he considers each dollar of loss separately. This approach by-passes a major stumbling block to utilization of small-risk experience. It also provides a much-needed theoretical explanation of the multi-split experience rating plans, in which this principle was first developed on a basis which, although intuitive, was nonetheless essentially sound.

The stumbling block to which reference has just been made is the variation between risks in the probability distribution of claims by size. If such variation is recognized, how can it possibly be measured when only a very small number of claims has been incurred by any but the largest risks? The measurement of credibility for each dollar of loss separately neatly substitutes for this apparently insoluble problem another that can be solved. By calculating the means and variances of the frequencies of successive increments to loss we arrive at a formula which gives high credibility to the first dollar of loss, lesser credibility to the fiftieth dollar, still less to the hundredth, and finally gives zero credibility to increments of loss in excess of some maximum value. By suitable gradation of credibility, an appropriate discounted value is provided for every possible size of loss.

The credibility of the t'th dollar of loss is given in the author's Eq. (4) as

$$(3) \qquad Z_t = \frac{Nm_t}{Nm_t + m_t^2/\sigma_t^2} = \frac{\underline{E}(f_t)}{\underline{E}(f_t) + m_t^2/\sigma_t^2}$$

in which $\underline{E}(f_t)$ is the expected number of claims for a risk with exposure equal to $\bar{N}$ units, while m_t and σ_t^2 are the mean and variance of the inherent hazard (as measured by the expected number of claims) per unit of exposure. The subscript t means that we are referring to the t'th dollar of loss, hence frequencies are counted only for claims of t dollars or more.

Although the author mentions a reference in which derivation for this equation is given, it is interesting to notice how immediately it follows from his first criterion for experience rating.

It should therefore be derivable from Eq. (2) of this review, in which the first criterion is mathematically reflected. The author's Eq. (4) can be written:

$$(4) \qquad Z_t = \frac{\sigma_t^2}{\sigma_t^2 + m_t/N}$$

The value σ_t^2 is the variance, within a class, of the expected number of claims per exposure unit, hence corresponds for unit exposure to the term σ_E^2 in Eq. (2) of this review. Under the Poisson assumption with respect to the probability distribution of the number of claims incurred by a given risk under constant hazard in a given period of time, the variance of the number of claims would be equal to the expected number, or m_t per unit of exposure. For N exposures the variance of the indicated pure premium per unit of exposure for the t'th dollar of coverage would be only m_t/N, however, which corresponds to the terms σ_A^2 in Eq. (2).

It is noteworthy that the author's Eq. (4) yields the same modification for the increment of rates provided for the t'th dollar of loss as does Hewitt's formula based on the negative binomial.[1] This is seen from the following:

If we count exposure in units of time, the variance between risks becomes $N^2\sigma_t^2$, and the variance of the number of claims sustained by one risk becomes Nm_t. We then have for N time units:

$$(5) \qquad Z_t = \frac{N^2\sigma_t^2}{N^2\sigma_t^2 + Nm_t} = \frac{\sigma_t^2}{\sigma_t^2 + m_t/N}$$

Where C claims have been incurred, as compared with Nm_t expected, the

[1] Charles C. Hewitt, Jr., "Negative Binomial Applied to the Canadian Merit Rating Plan for Individual Automobile Risks", P.C.A.S. XLVII, 1960.

modification applicable to that portion of the rate provided for the t'th dollar of loss becomes

$$(6) \qquad M_t = \frac{\sigma_t^2}{\sigma_t^2 + m_t/N} \cdot \frac{C}{Nm_t} + \frac{m_t/N}{\sigma_t^2 + m_t/N}$$

$$(6a) \qquad M_t = \frac{1}{m_t} \cdot \frac{C\sigma_t^2 + m_t^2}{N\sigma_t^2 + m_t}$$

Hewitt's formula (substituting N for his s) is:

$$(7) \qquad M = \frac{a}{r} \; \frac{r + C}{a + N}$$

where $a = m_t/\sigma_t^2$ and $r = m_t^2/\sigma_t^2$ in the notation used here. Equations (6a) and (7) are the same when these substitutions are made.

An example of a practical use of the formulas presented in this paper is provided by their application to a sample of 862 Homeowners risks studied by the writer. It was found that the calculated value of σ_t^2 turned "negative" for values of t in excess of premium. Negative values of σ^2 are, of course, impossible. The value of σ_t^2 was calculated by subtracting the expected number of claims, which equals the Poisson variance and corresponds to a homogeneous population, from the actual variance of the number of claims per risk. There will always be some value of t, however, for which a finite number of risks will generate not more than one claim apiece in a finite period of time, regardless of the variance between their means. The variance of such an observed distribution will of course be only binomial (npq with n equal to 1) hence less than Poisson (p). Until a more powerful method of analysis is recognized, it appears necessary to regard the value of t for which the observed value of σ_t^2 equals m_t as the limit beyond which further increments of loss have zero credibility.

As an interesting sidelight upon this study, it was found that when losses were discounted by means of a geometric progression similar to that underlying the multi-split plan used in Workmen's Compensation, the result was not significantly different from that obtained by applying the credibility formula derived in terms of the t'th dollar of loss. Such a simplification, of course, is most welcome in practical applications.

An important point is raised by the author to the effect that the parameters of an experience rating plan should be derived from experience. The need for doing so in connection with small risk experience rating, or merit rating, has long been recognized. This may have been because under merit rating plans a small number of classes can be set up to correspond to the several debit and credit groups established under such plans. For other experience rating, however, it would be no less appropriate to tabulate experience by the amount of the modification, and there is no real obstacle to arranging for this to be done. Such a study would provide a valuable check on the actuarial soundness of plans in current use, although it would not guarantee that they are the most efficient of possible plans.

In the latter part of his paper the author develops a method of experience rating which makes the least possible use of existing rates, consistent with a selected maximum degree of instability in collectible rates. The degree of instability is expressed in terms of the effect of a single loss. The conclusion reached—that up to the maximum single loss the credibility should be unity and that over this amount the credibility should be zero—is by no means intuitively obvious. The usefulness of this form of experience rating lies in those areas, such as new coverages or where a class is known to be very heterogeneous, where there is little confidence in the accuracy of established rates for individual risks or even, perhaps, overall.

An intermediate approach, not mentioned by the author, is provided where credibility is taken to be inversely proportional to the coefficient of variation of losses, thereby limiting but not minimizing the variance of formula rates. Under this procedure we say, in effect, that we don't know just how accurate the established rates are, but we do have a fair amount of confidence in them. We will therefore give as much credibility to the experience as we can, subject to a maximum variance in formula rates (as Z approaches zero) equal to the sum of the variance of established rates plus the variance which corresponds to full credibility. For intermediate credibilities we will accept a variance in formula rates equal to the sum of the variance corresponding to full credibility plus the product of the square of the complement of the credibility times the variance of established rates.

An incidental point, mentioned by the author in connection with the approximation of claim distributions, deserves comment. He suggests that some available data indicate that the log-normal curve is appropriate for fitting claim distributions in fire insurance as well as in the casualty lines. The writer has found this to be the case except as the policy limit is approached. In that region a graph on log-normal probability paper curves upward. Available data are insufficient to show precisely how the function behaves in the immediate vicinity of the policy limit, but it seems reasonable to believe that a discontinuity exists at that value because of the probability of total losses.

In conclusion, I should like to commend the author for having contributed one of the most scholarly and valuable papers to be found in our Proceedings.