# ON THE CREDIBILITY OF THE PURE PREMIUM

ALLEN L. MAYERSON, DONALD A. JONES,
NEWTON L. BOWERS, JR.

VOLUME LV, PAGE 175

## DISCUSSION BY JEFFREY T. LANGE

"It is the purpose of art to give a clearer picture of reality."
— Piet Mondrian

One of the purposes of constructing a model or theoretical abstract of an actual situation is to be able to see it more clearly. Many ratemakers have recognized practical deficiencies in credibility tables. Rates for low volume lines of insurance (and for low volume territories of major lines) are sometimes increased one year only to be reduced in the next year. They have also noted that the formal credibility standards are insensitive to the line of insurance, in that the same credibility factors are assigned in spite of differences in the distribution of number of claims and claim amounts. In their theoretical paper, Messrs. Mayerson, Jones, and Bowers have provided the ratemakers with a clear picture of the weaknesses of existing procedures and a practical tool for recomputing credibility standards.

It would be hard to disagree with the authors' conclusion that the existing credibility standards for automobile and general liability lines are too low. Assuming a Poisson distribution for the number of claims and using the 1963 size of claim data published by the National Bureau of Casualty Underwriters and the Mutual Insurance Rating Bureau, the reviewer obtained a 100% credibility criterion of 3,434 for automobile (private passenger) property damage liability data with a probability of 90% that actual losses will be within 5% of expected losses. This solution compares with the authors' 4713 and the current 1084. Calculations for general liability coverages also confirmed the authors' conclusion.

The authors' second and third conclusions appear equally valid. The impact of recognizing the positive skewness of the claim amount distribution and the use of the negative binomial in place of the Poisson would further raise the credibility criteria. However, as noted by the authors and by Mr. Nelson, in his discussion of this paper, the numerical significance of

these two refinements is much less than that of reflecting the size of claim distribution.

The fourth conclusion that credibility criteria vary by coverage can be illustrated with the following credibility criteria computed from the 1963 automobile data cited above and from 1967 general liability data published by the Insurance Rating Board and the Mutual Bureau.[1] (In the following paragraphs, full credibility is defined as being within 5% of the expected pure premium with a probability of .90.) While a credibility standard of 3434 claims was indicated for private passenger property damage liability coverage (total limits), the bodily injury coverage would require 3931 claims when losses are limited to $5000 and 5098 claims when losses are limited to $10,000. A later sample of claims produced a 6241 claim standard for unlimited (or total limits) bodily injury private passenger. Thus, within the private passenger subline, there was a substantial difference between the bodily injury and property damage coverages (6241 versus 3434) and the limitation of losses to $10,000 and $5000 resulted in substantially reduced credibility requirements.[2]

The use of private passenger data for New York only gave credibility standards about 20% below countrywide. Countrywide commercial car data (as opposed to private passenger) yielded bodily injury results almost identical to the private passenger while the property damage requirement was 4462 versus 3434.

For the general liability coverages, even greater variation was observed. For owners, landlords and tenants insurance, 100% credibility points of 4583 and 11,881 were obtained for bodily injury and property damage respectively, while for manufacturers and contractors bodily injury coverage, a 3672 criterion was obtained.

It should be noted that the sample sizes used in these calculations are substantial. For example, for private passenger cars over 300,000 claims were used while the O.L.&T. bodily injury sample was more than 75,000. The substantial differences coupled with the ease of making the calculations suggest that, in the future, credibility standards, being more or less tailor-made to the situation, might be subject to much more variation by line and state.

---

[1] The author wishes to acknowledge the assistance of J. Robert Hunter in the preparation of this section.

[2] A similar conclusion was reached by L. H. Roberts in "Credibility of 10/20 Experience as Compared with 5/10 Experience," *PCAS* Volume XLVI, p. 235.

"On the Credibility of the Pure Premium" leaves unresolved several points concerning credibility. First, what rule should be used for partial credibility? Mayerson[3] rejected $Z = (n/N_{1.00})^{1/2}$ (the square root rule) in favor of $Z = \frac{n}{n+k}$, but this latter formula is inconsistent with having a 100% credibility point as derived in this paper. Such formulas for partial credibility were rejected by Maguire[4] in his 1969 paper on credibility. Second, is the authors' approach to credibility, based on Perryman, the proper one? Would other approaches — Maguire, Buhlmann,[5] or Braverman[6] — yield a "better" result? Third, if one accepts the authors' approach, how does one select the values of $P$ and $k$. In the case of the model leading to 1084 claims, whether one selects $P$ at .90, .95 or .99 (all reasonable) and $k$ at .01, .05 or .10 (also reasonable) leads to a variation of 100% credibility criteria from 271 to 10,623. This range is more significant than the difference between the negative binomial and Poisson distributions or even between the old credibility standards and the indicated. Thus, the authors have not given (and did not claim to give) a formula which yields a unique 100% credibility point suitable for all calculations.

As far as the classical theory has been developed, given a particular loss distribution (for a line and state) one may obtain a unique criterion for 100% credibility once an arbitrary $P$ and $k$ are selected. In some practical situations one does not seek a unique criterion. For example, in a rate filing, it would be desirable to use two or more criteria. For determination of territory relativities, the disutility of giving too little belief to the latest indication may mean the loss of the opportunity to write additional business, or the loss of existing business, in certain territories (i.e. the loss of the opportunity of a profit on certain units). Since the relativity procedure is balanced, giving too little credibility does not result in a loss in premium revenue on all units but rather charges that are too high for some units and too low for others.

In the calculation of statewide rate level, on the other hand, the indication is credibility weighted with no change, and any reduction in credibility results in a reduction in premium revenue on all units, since in an inflation-

[3] Mayerson, A. L., "A Bayesian View of Credibility," *PCAS* Volume LI, p. 85.
[4] Maguire, R. D., "An Empirical Approach to the Determination of Credibility Factors," presented at the Spring, 1969 Meeting of the Society of Actuaries.
[5] An explanation of Buhlmann's approach to credibility is given in C. C. Hewitt's discussion of this paper.
[6] Braverman, J. D., "A Critique of Credibility Tables," *Journal of Risk and Insurance* Volume 34, p. 409.

ary economy increases are more common than decreases. The disutility of too little credibility is much greater in the case of statewide rate level than in the case of territory relativity calculations, since the former can lead to a dollar loss on all units while the latter results only in the lost opportunity for profit on some units. In other words, the mathematical and statistical tools cannot be separated from the decision procedure — the business situation — itself. If one accepts this premise, then the selection of $P$ and $k$ becomes an exercise in statistical decision theory once the disutilities of the various outcomes are postulated. This latter step is, of course, a subjective one and can be performed only if one accepts the subjective, or Bayesian, view of probability. Apparently, the classical Neyman-Pearson solution of credibility problems is only of use to Bayesian statisticians.

## DISCUSSION BY JOHN S. McGUINNESS

This is a fundamental paper of great significance. Besides the concrete advance it explicitly reports, it also has several substantial implicit qualities. These are easy to miss but they are worthy of specific recognition for the valuable instruction they provide.

Primarily, the paper provides a sound and distribution-free theoretical basis that permits the development of a specific criterion (a precise number of claims) for ascribing full credibility to data that reflect both relative claim frequency and average claim cost for a single-parameter class of risks. It appears to provide an objective basis for overcoming two present and contrasting deficiencies in ratemaking practice. The first deficiency is the current neglect of the average claim cost element (and its variability) in determining a class credibility, and the second is the present common neglect of the relative frequency element (and its variability) in making time-series adjustments.

The author's approach, in using the total amount of claims (T) as the major variable of which the variation is measured, is a less easily visible quality. It is a clever and rewarding change in viewpoint. By focusing their attention on this aggregate or collective figure, rather than on the much smaller pure premium, they have for their analysis a statistically much more manageable datum and one for which it is much easier to determine a mean and an objective measure of variability. This difference in approach appears to parallel precisely the difference between approaching