REPORT

AN INTRODUCTION TO CREDIBILITY THEORY

BY

L. H. LONGLEY-COOK

1. PREFACE

Credibility Theory is one of the cornerstones of actuarial science as applied to casualty and property insurance. Although the literature of this theory is extensive, there is no elementary introduction at present available. Nearly all the numerous papers[1] bearing on credibility theory which have appeared in the Proceedings of the Casualty Actuarial Society are difficult to follow without a knowledge of the subject and many of them are very long and involve fairly abstruse mathematics. At the request of the Educational Committee of the Society, the author has prepared this introduction to provide actuaries and others interested in credibility theory with a framework into which they can fit these papers so that they can better appreciate the large body of research which has been carried out in this field. The author has tried to keep mathematics to a minimum and has concentrated on principles rather than details. It must be stressed that this is merely an introduction to the subject and it is only by studying the many original papers that a full comprehension of the subject can be obtained.

2. MEANING OF CREDIBILITY

The basis for these credibility formulas has been a profound mystery to most people who have come into contact with them.

—Arthur L. Bailey[2]

The word credibility was originally introduced into actuarial science as a measure of the credence that the actuary believes should be attached to a particular body of experience for rate making purposes. Thus we say that the loss experience under a new class of insurance is "still too small to be credible", implying that the experience which will develop in the future may well be very different from that so far collected, and also implying that we have more confidence in our prior knowledge based on other data such as current rates for similar classes. Again, the statement that the private passenger automobile liability experience in Pennsylvania is "fully credible for rate making", implies that the experience, after adjustment by trend factors, is adequate to establish the overall rate level in the state without reference to previous rates or data or to experience in other states.

In many cases a body of data is too small to be fully credible but large

[1] See Appendix A.

[2] Credibility Procedures—CAS XXXVII, p. 7 (1950)

enough to have some credibility. A scale of credibility has been established which gives 0 credibility to data too small to be any use for rate making and 1 credibility to data which are fully credible. Credibility theory is concerned with establishing measures of credibility and standards of full credibility.

Arthur L. Bailey[2] has pointed out the special recognition given to prior knowledge in credibility theory

> "At present, practically all methods of statistical estimation appearing in textbooks on statistical methods or taught in American universities are based on an equivalent to the assumption that any and all collateral information or *a priori* knowledge is worthless. . . . It appears to be only in the actuarial field that there has been an organized revolt against discarding all prior knowledge when an estimate is to be made using newly acquired data."

As a corollary of the recognition of prior knowledge, the amount of credibility to be attached to a given body of data is not entirely an intrinsic property of the data. For example, there is always stated or implied in any measure of credibility the purpose to which data are to be used. Thus certain data obtained from the reports of Fire Marshalls in the State of Oregon may have a high credibility for establishing the variations of fire rates by protection grading in that state. The same data will have a lower credibility when applied to establishing the variation of fire rates by protection grading in the states of Pennsylvania or New York.

The term credibility has been extended to represent the weight to be given to certain data in various experience rating formulae. Much confusion has resulted from assuming that the credibility of data used in an experience rating plan is the same as the credibility of the same data if it were to be used for some other purpose such as independent rate making.

Hence, we see that credibility is not a simple property of data which can be calculated by some mathematical formula as can the standard deviation or other measures of the effect of chance variation on a body of statistical data. While credibility and statistical variance are related, the former is meaningful only against a stated or implied background of the purpose for which the data are to be used and a consideration of the value of the prior knowledge available.

3. THE NEED FOR A MATHEMATICAL MODEL

> *The calculus of probability is a fascinating subject, and one which is destined to play a large part in actuarial science; and a day may come when it can truly be said of the actuary that he has fused together the theories of finance and probability.*
>
> —E. W. Phillips[3]

The application of mathematics to science follows a fairly standard pattern. First certain "laws" are established usually by a combination of careful research and general reasoning. From these laws a mathematical model is constructed and is tested against actuality. The model is then used for further research and as a means of forecasting what will occur in designated

[3] Biometry of the Measurement of Mortality—privately published, p. 5 (1935)

special circumstances. In the "exact" sciences the models are very close to reality; in the less exact sciences the models are only approximate. Since the model follows directly from the laws by the application of pure mathematics, it is convenient to use the term model for both the laws themselves and the mathematical development therefrom.

Thus in dynamics the laws are usually Newton's laws of motion and the model deduced therefrom allows, for example, the accurate prediction of an eclipse of the sun many years hence. This model is almost, but not completely, exact.

In the actuarial science of life insurance the original law was a rate of mortality which depended upon the age and sex of the life but was independent of any other variable. The actual rates of mortality would depend upon the body of lives under review but for any individual problem the law was assumed to hold. The model developed from this law was the mortality table and the whole theory of life contingencies, without which it would not be possible to transact life insurance as we understand it today, is based thereon. The model was later developed to allow for selection and further developments were required to deal with special problems such as impaired lives and options. However for much actuarial work in the life insurance field the original model is still the basis of all calculations. It was early realized[4] that the rate of mortality represented the average mortality of all lives in the group only, but it was rarely if ever necessary to reflect this in the actuarial work of life insurance and nearly all life insurance calculations can be made on assumption that the rate of mortality applies to each individual life.

In other fields of insurance the development of suitable models has been difficult and the applications of the models nearly always require not only the use of average values, but a consideration of the distribution of the variations from the average. In the United States the development of these models and their application to practical problems have been associated mainly with rate making and credibility. While the word credibility was originally introduced to indicate the credence that the actuary believes should be attached to a particular body of experience for rate making purposes, the use of the term has been extended to many rate making techniques associated with this general idea. On the continent of Europe, the development and applications of these models have been referred to as the Theory of Risk and the main application has been the study of the effects of chance variation on the surplus of an insurance company.

4. STATISTICS FOR INSURANCE RATE MAKING

We who serve our Lady Casualty,
Should be of all men first,
Most resolutely to hope for the best,
Most wisely prepare for the worst.

—Clarence W. Hobbs[5]

Before we can begin to construct a model appropriate for the study of

[4] On the Improvement of Life Contingency Calculation—E. J. Farren. Journal of the Institute of Actuaries, Vol. 5, p. 185 (1855)

[5] The Lady Casualty and Her Servitors—CAS XXVI, p. 168 (1935)

problems of casualty and property insurance we must understand clearly the purpose for which the model is primarily intended and certain of the characteristics of this branch of insurance. It is important to realize that casualty insurance is a contract of indemnity and hence the amount payable in case of a loss must be determined by the individual circumstances of each case and may depend upon a legal action and a jury verdict. For this reason loss experience will never be stable for any length of time and rate revisions are frequent. We are concerned with analyzing the past, mainly to enable us to develop premium rates for the future.

It is appropriate here to quote some remarks of Arthur L. Bailey[6], the actuary who has contributed more than anyone else to our knowledge of credibility, on the difference in philosophy of the casualty actuary and the statistician in the more usual fields of statistical study.

> "First, there is the belief of casualty underwriters that they are not devoid of knowledge before they have acquired any statistics. This belief is probably held by operating personnel in all businesses. When a new form of insurance is initiated or a new classification or territory established, there may be a considerable variety in the opinions of individual underwriters as to what the rate should be; but the consensus of opinion invariably produces a rate. This rate soon becomes embedded in the minds of the underwriters as the 'right' rate. Later, when statistics as to the actual losses under the new coverage, classification or territory, finally are acquired, the problem is not 'what should the rate have been?' but 'how much should the existing rate be changed as a result of the facts observed?' In revisions of rates for regular coverages, classes and territories, this is always the question.
>
> "The statistical methods, developed by the mathematicians and available in the standard textbooks on statistical procedures, deal with the evaluation of the indications of a group of observations, but under the tacit or implicit assumption that no knowledge existed prior to the making of those particular observations. The credibility procedures, used in the revisions of casualty rates, have been developed by casualty actuaries to give consistent weightings to additional knowledge in its combination with already existing knowledge.
>
> "A second belief of casualty actuaries is that they are in a continuing business. Also that a more or less wide spread of risks is being taken at any time. The rate maker in such an organization as the National Bureau of Casualty Underwriters literally has thousands of rates to be revised at relatively frequent intervals. Being called upon to make a large number of estimates, the casualty statisticians can relinquish the condition, usually imposed by other statisticians, that each estimate be unbiased. In its place they may impose the less restrictive condition that a particular group of estimates be unbiased in the aggregate. This permits them to make a material reduction in the error variances below what could be obtained by applying the usually taught and presented methods of statistical estimation. It produces another type of credibility formula which appears to be unique to casualty insurance.

---

[6] Discussion by Arthur L. Bailey, Journal of the American Teachers of Insurance, Vol. 17, p. 24 (1950)

"The third peculiarity is that casualty underwriters consider each insured to differ from all other insureds. For example, each automobile driver is assumed to have habits and eccentricities unlike any other; each fleet of trucks is assumed to travel routes and engage in operations which make its hazards different from all others, even those engaged in the same industry in the same territory. The propriety of this assumption has been verified in so many instances that the differences between risks has become a basic concept or axiom. Experience rating plans are used in almost all lines of casualty insurance to measure the peculiarities of individual risks.

"Despite this uniqueness of the 'inherent hazard' of different insureds, each and all of them are subject to the vagaries of chance and the random errors of classification and measurement common to all statistics. Statistical methods generally taught and published in textbooks deal with populations for which the entire variation is produced by the vagaries of chance or the random errors of measurement. Populations in casualty insurance, however, consist of individuals having a variation of expectations other than that due to these two items. Their inherent hazards must be assumed to differ even if it is impossible to postulate or to precisely measure the differences. This dealing with heterogeneous populations produces some very interesting results which most statisticians would sneer at as 'impossible', but which are, nevertheless, wholly sound and justifiable."

While the above remarks by Arthur L. Bailey are specifically directed to casualty insurance, similar considerations apply to property insurance.

5. DISTRIBUTION OF NUMBER OF LOSSES—THE FIRST MODEL

In the early history of navigation, we find it taken almost as the basis of the science that the compass needle pointed in a fixed direction, and that such a direction was due north. The utility of so simple an assumption in early days can scarcely be overrated. —E. J. Farren[4]

The scientist does not need to invoke Einstein's relativity theory for each simple calculation on moving bodies. In most cases the laws of motion developed by Newton are sufficient for his purpose. Nor need the actuary assume a complex model when a simple model will suffice. In most types of casualty and property insurance, more than one loss (accident) can occur in a year. In our first model it will be assumed that the probability of an accident in any period of time is the same for each individual exposure (person, automobile, etc.) and that it is proportional to the length of the time exposed. Further, it will be assumed that we are either concerned with studying the number of accidents and not the amounts of loss or, alternatively, that the amount of loss for each accident is the same.

If the number of exposures over a period of one year is n and the probable number of accidents in any period dt is nqdt, then the most probable total number of accidents in the year is nq and the average number of accidents to any one individual is q. It is necessary to determine the probability that an individual has exactly none, one, two, etc., accidents in the year and that the total number of accidents is within k percent of the most probable nq. The solution of this problem will be found in most textbooks on statistics since this is the well-known *Poisson Distribution* and is developed in Appendix B. The results are as follows:

The chance of exactly r accidents out of n individuals is

$$\frac{(nq)^r e^{-nq}}{r!}$$

In the case of a single individual this becomes

$$\frac{q^r e^{-q}}{r!}$$

The mean number of accidents is nq and the variance is also nq.

Further, the probability P that the number of accidents will be within $\pm$ 100k% of the expected nq is equal to

$$P = 2\left[\frac{1}{2\pi}\int_0^{k\sqrt{nq}} e^{-\frac{t^2}{2}}\,dt\right]$$

when n is large and k is not large. (See Appendix B).

This formula is used to establish standards of full credibility.

6. NUMBER OF CLAIMS REQUIRED FOR FULL CREDIBILITY

A dependable pure premium is one for which the probability is high, that it does not differ from the true pure premium by more than an arbitrary limit.

—Albert H. Mowbray[7]

When the actuary says certain data are fully credible, he is not implying that, if it were possible to collect another body of data of the same size under identical conditions the result would be for all practical purposes identical, but rather that the volume of data is adequate for rate making without reference to other experience data and without reference to the premium rates previously charged. While it may be of some interest to know what volume of data is required to meet the former test, it will normally be so large that it could never be available, and hence the enquiry is academic. Insurance data are unlike data available in biometric and similar studies where practically any volume of data desired can be collected if we go to the necessary trouble and expense. Normally we are concerned with the whole of the data for a particular classification and no further identical data are available. It is only when the data for a single insurance company are being studied that a larger volume of data, that of all similarly operated companies, may be available. Even in this case it is important to stress the words "similarly operated". In

[7] How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium? CAS I, p. 24 (1914)

many lines of business the experience of stock and mutual companies are not the same nor that of direct writers and agency companies and a combination of the experience of differently operated companies will not produce greater credibility for establishing the over-all rate level. However, it may well produce greater credibility for establishing rate differentials for classification sub-groups.

Our first model can be used to establish the number of claims required to reduce to negligible proportions the probable departure from the number observed which could be attributed to chance variation. On the basis of the formula set out at the end of the preceding section, using published statistical tables, we establish the number of claims necessary to meet some typical values of P (the probability that the number of accidents will be within $\pm$ 100k% of the expected number of accidents) and k as follows:

Maximum Departure from expected (k)	*Probability of meeting test (P)*		
	99%	*95%*	*90%*
	Number of claims required		
2½%	10,623	6,147	4,326
5 %	2,656	1,537	1,082
7½%	1,180	683	481
10 %	664	384	271

The figure corresponding to P equals 90%; and k equals 5%, namely 1,082, is frequently used as an accepted standard of credibility.

It will readily be appreciated that when a more realistic model is used and allowance is made for the variation in the amount of claim from accident to accident, the volume of data required to meet full credibility for, say, pure premiums, will be higher. (The development of the relationships of the credibility for Claim Frequencies, Claim Costs, Pure Premiums and Trends is set forth in Appendix C.) However, since the choice of P and k are in any case arbitrary, we can justify the use for Pure Premiums of the standard already established for loss frequency by a suitable variation in the values chosen for these factors. Further, in practice, the selection of too high a standard for full credibility would considerably delay the response of premiums to changed accident conditions and might well lead to over-all inadequacy of premium levels. The standard of full credibility is not normally important in itself, but is important as a means of introducing consistency in the rate making procedure and establishing proper relationships as respects reliability between different volumes of experience.

While the number of claims is the most appropriate measure for establishing credibility, it is not always the most practical one and it is often necessary to use premiums instead. A standard of $5,000,000 of premiums or some similar figure for full credibility has been customary in fire insurance although the volume should vary according to the average size of loss. A larger volume is theoretically necessary for industrial than for habitational risks.

The earliest paper in the Proceedings of the Casualty Actuarial Society on the standard for full credibility was presented by Albert H. Mowbray in

1914[7] and the most useful general reference to the subject will be found in a paper by Francis S. Perryman[8].

7. PARTIAL CREDIBILITY—RATE REVISIONS

How much weight the indications of specific volumes of data are to be given in the casualty business has continued to be a matter of individual judgment.

—Arthur L. Bailey[2]

If we have a certain loss frequency based on past experience and a new set of data which is not large enough to provide full credibility, on the basis we have accepted for the type of business under discussion, how should the two sets of information be combined for future rate making? Obviously we should be wrong to discard the new data because it is not fully credible, nor should we ignore the old rate which may be based on a vast volume of data. Clearly some combination of the two is required. It is tempting to proceed as follows. The probability of a claim in the old data is $p_0 = l_0/E_0$ where l_0 is the number of claims and E_0 the exposure units in the old data. The probability of a claim in the new data is $p_1 = l_1/E_1$ where l_1 and E_1 are the claims and exposure units in the new data. Combining the data from the two sources we have as the best estimate of the true probability

$$p = \frac{l_0 + l_1}{E_0 + E_1}$$

There are a number of reasons why this procedure is not practical. We rarely know the precise basis of p_0. It may be based on a certain number of claims and exposure units. It may be adjusted to reflect in part some previous experience, or it may reflect informed judgment. Further it may be partially obsolete because of trends which have occurred in the interval since the data were collected. It must also be noted that a formula of this form does not produce $p = p_1$ if p_1 has full credibility. Another approach is therefore necessary.

If we have a sufficient volume of data for p_1 to meet our requirement of full credibility then

$$p = p_1$$

Again if the volume of data is so small as to be meaningless it is probably wise to assume

$$p = p_0$$

For all other volumes of data, p must lie between p_0 or p_1. Expressed mathematically this means

$$p = p_0\,(1\text{-}Z) + p_1 Z$$

or

$$p = p_0 + Z\,(p_1 - p_0)$$

where Z lies between 0 and 1. It will be noted that if the data have full credibility $Z = 1$ and if the data have no credibility $Z = 0$. The value of Z is called the credibility assigned to the new data.

What value of Z, credibility, should be assigned to a volume of new data less than that to which full credibility is assigned? To get some insight into

[8] Some Notes on Credibility—CAS XIX, p. 65 (1932)

this problem we will first turn to the approach suggested at the beginning of this section. The expression for p there developed may be written

$$p = \frac{E_0 p_0 + E_1 p_1}{E_0 + E_1}$$

$$= \frac{(E_0 + E_1) p_0 + E_1 (p_1 - p_0)}{E_0 + E_1}$$

$$= p_0 + \frac{E_1}{E_0 + E_1} (p_1 - p_0)$$

Various assumptions can be made as to the relationship of E_0 to the value of E_1 required for full credibility. However, whatever assumptions we make we shall get a somewhat similar curve which should give us insight into the true shape of the Credibility curve. If we assume E_1 is equal to twice E_0 when E_1 has a volume corresponding to our criterion of full credibility we get the following results.

Volume of new data as a percentage of data for full credibility $E_1/2E_0$	*Indicated Credibility* $E_1/(E_0 + E_1)$	*Column (2) multiplied by 1.5*
100%	.67	1.00
90%	.64	.96
80%	.62	.92
70%	.58	.87
60%	.55	.82
50%	.50	.75
40%	.44	.67
30%	.38	.56
20%	.29	.43
10%	.17	.25
0%	0	0

The final column is the suggested scale of credibility obtained by increasing all the indicated credibilities in the same proportion so as to make the indicated credibility 1.00 for the volume of data we have agreed will correspond with full credibility. Other assumptions can be made concerning the relationship of E_1 to E_0. If we assume full credibility corresponds to a lower multiple of E_0 the curve becomes more nearly a straight line. If we assume full credibility corresponds to a higher multiple of E_0 the curve becomes less steep at its upper end and steeper at its lower end. It will be noted that the "indicated credibility" takes the form

$$Z = \frac{n}{n + k}$$

when n is the number of losses or some other measure of the volume of data and k is a constant. This is a form originally suggested by Albert W. Whitney[9].

For another approach to the problem we can take a quotation from a

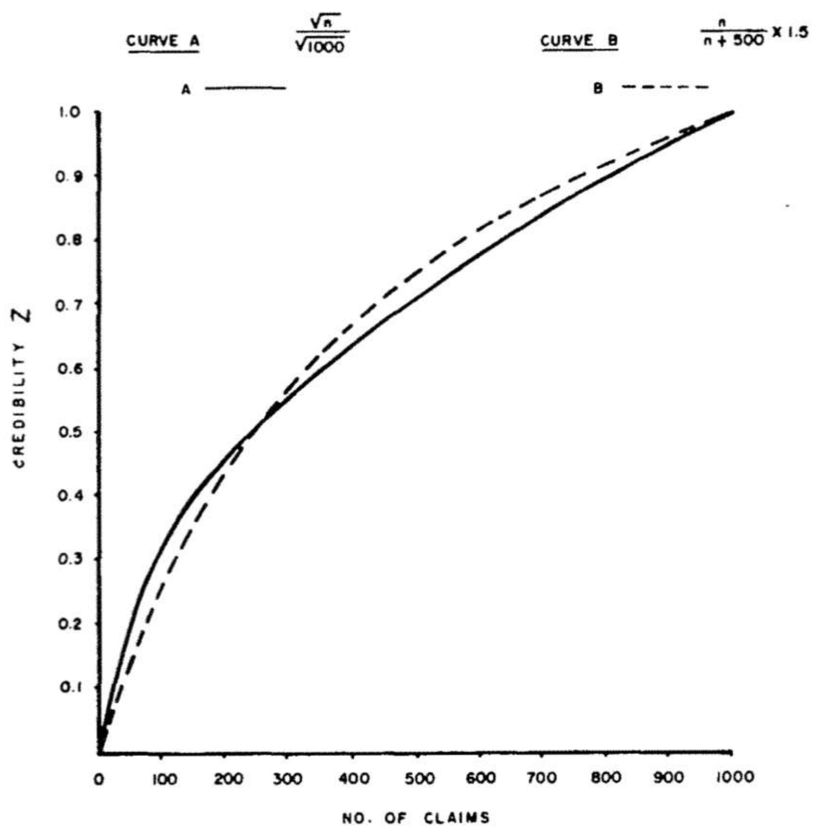
[9]The Theory of Experience Rating—CAS IV, p. 274 (1918)

paper by Francis S. Perryman[5]. "The reasons prompting the use of this do not appear very explicitly in casualty actuarial literature, but it seems to be based on the rule used in 'combination of observations' (in such sciences as astronomy, engineering) that the best weight to be assigned an observation is the reciprocal of its standard deviation: according to this the relative weights of two experiences, one (exposure n) entitled to 100% credibility and other (exposure $\frac{n}{r}$) would be in the ratios of the reciprocals of their standard deviations or as $\frac{\sqrt{n}}{\sigma}$ to $\frac{\sqrt{\frac{n}{r}}}{\sigma}$ that is 1 to $\frac{1}{\sqrt{r}}$.

"The rule seems plausible and practical. It is to be noted, however, that the principles upon which it was derived for use in other branches of science are not especially applicable to casualty rate making."

Values obtained by this method are very similar to those developed in the previous approach as the following table, based on 1000 claims for full credibility, shows.

GRAPHICAL REPRESENTATION OF CREDIBILITY CURVES



	Credibility	
No. of Claims	$\frac{\sqrt{n}}{\sqrt{1000}}$	$\frac{n}{n+500} \times 1.5$
1000	1.00	1.00
900	.95	.96
800	.89	.92
700	.84	.87
600	.77	.82
500	.71	.75
400	.63	.67
300	.55	.56
200	.45	.43
100	.32	.25
0	0	0

The $\sqrt{n}$ approach is the one most generally used at the present time.

When we come to consider a less simple mathematical model and take note of the distribution of losses by size, we shall find that the distribution of credibility by volume of business will not necessarily follow the simple $\sqrt{n}$ rule.[10]

Another approach to credibility which has been tried in fire insurance is to measure credibility by the average annual variance of the ratio of incurred losses to earned premiums from the five year average.[11] The loss ratio should be first adjusted for trend, as indicated by the all classifications' loss ratios, and also for rate revisions; but even without these adjustments, a good idea of credibility can be obtained by noting the amount of stability in loss ratios from year to year.

In addition to the papers referred to above the problem of partial credibility has been studied extensively by Arthur L. Bailey[2, 12, 13].

8. STABILITY IN RATE REVISIONS

> *A rate-level determination upon statisical foundations is always a compromise between the two conflicting considerations of responsiveness to recent experience indications and stability sufficient to avoid frequent and undue disturbances in the field.*
>
> —T. O. Carlson and L. H. Longley-Cook[14]

We must always remember insurance is a business and rates and premiums are more than mere statistical developments. They determine the actual sums

[10] Robert L. Hurley, A Credibility Framework for Gauging Fire Classification Experience—CAS XLI, p. 161 (1954)

[11] Experience Credibility Formula—Middle Department Association of Fire Underwriters (1949)

[12] Sampling Theory in Casualty Insurance—CAS XXIX, p. 50 (1942) and CAS XXX, p. 31 (1943).

[13] A Generalized Theory of Credibility—CAS XXXII, p. 13 (1945)

[14] Multiple Line Insurance—Michelbacher p. 98, McGraw-Hill (1957)

of money payable for insurance coverage and these sums may be considerable. While it is essential that premium rates correctly follow overall trends, year to year fluctuations in rates can prove most unfortunate. Such fluctuations not only cast doubt in the mind of the public and of regulatory authorities on the correctness of the rate making procedure, but may have a number of side effects, such as leading to the cancellation and rewriting of a number of policies prior to expiration.

The method of establishing partial credibility in Section 7, while having a reasonable plausibility and the sanction of long practical use, is open to quite serious criticism. Where the volume of current data is small, the volume of past data on which the old rates were based will almost invariably be small also and hence the relative weights to be given to old and new data should be much the same whatever the volume of data. If we are concerned with revising a rate for some special class of insurance which is in no way related to any other class, the partial credibilities developed in the previous section are probably too small and higher credibility factors are almost certainly desirable. However, such cases rarely occur in practice and, as we shall see in the next section, individual rates are normally part of a pattern of rates which cannot be ignored at a rate revision. It is mainly for this reason that the system of partial credibilities has proved satisfactory in practice.

The actuary will normally seek, by suitable grouping, to establish a body of associated data which is sufficiently large to be fully credible for determining the overall rate level. Thus he will group together all private passenger automobile liability insurance in the State of Pennsylvania, all fire insurance in California, all personal property floater business nationwide, etc. Subgroups within these broad groups will not normally be fully credible and the partial credibility techniques discussed in the previous section will be used for rate revisions. The indicated rate revisions so developed will usually be adjusted by rule or by judgment to avoid major variations in individual rates. Unless the overall data indicate a major change in rate levels, it is often the rule to limit individual revisions to increases or decreases of 25%.

9. RATE RELATIVITIES

In the several papers in our Proceedings and in the Transactions of the Actuarial Society of America dealing with compensation premium or rate making, the starting point has been a classification pure premium derived by the well-known formula, $\pi = L/P$. It has been generally recognized that it will be impossible to determine the pure premiums in this way for each classification, and that some process of association must be resorted to in order to develop premiums for those classifications where the data is insufficient.

—Albert H. Mowbray[15]

When we come to consider the revision of any important body of rates such as Fire insurance in New York or Automobile Liability insurance in

[15] The Determination of Pure Premiums for Minor Classifications on which the Experience Data is Insufficient for Direct Estimate—CAS II, p. 124 (1915)

Pennsylvania, we normally have a volume of data which is fully credible for the purpose of determining the overall rate level which is required. However, each of these major classes includes a very large number of subclasses and there will never be sufficient data to provide credible rate revisions for each individual rate. Most rating systems contain patterns of association between various rates to which the term "rate relativities" is applied. Thus, for each sub-territory of an individual state, there will be a relationship between the liability rates for various driver classifications which will follow a definite pattern. The determination of these patterns is one of the most important parts of rate making. In, for example, the complex problem of schedule rated fire insurance risks, the various relativities depending upon types of construction are based almost entirely upon engineering judgment; in less complex rating plans, the assumption is made that the rates in one classification are related by some simple rule to the corresponding rates in another classification. The rule, which may be based on judgment alone or may be derived by some special study, will usually take the form of a percentage or constant differential but more complex rules are occasionally used. To make this method clear let us suppose that the rates for a certain type of insurance within a certain state have three types of classification: (1) territorial, (2) type of insured, (3) degree of inherent hazard. If there are five breakdowns in each classification—five territories, five types of insured, etc.—there will be $5 \times 5 \times 5 = 125$ subgroups of the data, no one of which could be individually credible. If however we assume that the change in rates between territories raises or lowers all rates in equal proportion, then by grouping all data by territory only with the necessary adjustments to reflect differences in distribution by the other classifications we can establish territorial differentials with a reasonable degree of credibility. By regrouping the data in other ways, other classification differentials can be established. We may liken our statistics to a large crumbly loaf cake, which we may cut in slices to obtain easily edible helpings. The method of slicing may be chosen in different ways—across the cake, lengthwise down the cake, or even in horizontal slices—but only one method of slicing may be used at a time. If we try to slice the cake more than one way at a time, we shall be left with a useless collection of crumbs.

It has been pointed out earlier that the combination of stock and mutual data will not normally provide a more credible estimate of the overall rate level because the experience of the two groups may be fundamentally different; however, such a combination and other similar combinations will usually provide more credibility in establishing rate relativities between subclasses.

Recently Robert A. Bailey and LeRoy J. Simon[16] have suggested that with modern electronic computors it should be possible to determine classification rate relativities in a single procedure giving the correct credibility weight to each subdivision of the data. This development would be a valuable advance in rate making techniques if we could introduce into the input not only the total data but also certain judgment rules which would insure an orderly pattern in the rate relativities.

It will be noted that in rate revisions we try to avoid as much as possible the use of data with only partial credibility and thus keep the rates as re-

---

[16] Two Studies in Automobile Ratemaking—CAS XLVII, p. 1 (1960)

sponsive as possible to the latest data. Where this is not possible, we must use credibility techniques and judgment to maintain a stable rate structure.

10. DISTRIBUTION OF LOSSES BY SIZE—THE SECOND MODEL

> *As the pure premium is the accident frequency multiplied by the average claim cost, we must see how possible variations in the average claim cost affect the pure premium and how we must modify our credibility requirements accordingly.*
> —Francis S. Perryman[8]

Arthur L. Bailey[12] in a discussion of the distribution of losses by size writes,

> "The various distributions of claims by size of claim are uniform in that they all exhibit a concentration of frequency at the low amounts with a tapering off of the frequencies up to and including very high amounts. This produces a skewness far in excess of that usually encountered in the study of frequency distributions. The only type of frequency distribution which has been found to fit these distributions of claims by size is the Normal Logarithmic Distribution. Tests of the goodness of fit of this type of distribution have indicated that, except for the concentration of claims at such round-figures values as $50, $100, $500 and $1,000, the departures of the actual distributions from the Normal Logarithmic are not greater than would frequently occur in samples of the size tested."

It may be explained that the Normal Logarithmic or Log-normal distribution implies that, if a curve is plotted of the distribution of the logarithms of the amounts of individual claims, a normal curve will result. Arthur L. Bailey in the same context goes on to state,

> "The only condition necessary to produce a Normal Logarithmic Distribution is that the amount of an observed value be the product of a large number of factors, each of which is independent of the size of any other factor. Reflection as to the conditions entering into the determination of the amount of a claim settlement in casualty insurance, the variations in the seriousness of accidents for which claims are made, and all of the factors eventually recognized in making the final settlement makes it apparent that the necessary condition is at least approximated in the data with which we are concerned. When this condition is met, the logarithms of the observations become the sum of a large number of independent elements, which is the only condition necessary to result in a Normal Distribution. Thus, we shall expect to find the logarithms of the claim amounts normally distributed."

The generalized Normal Logarithmic Distribution provides an additional degree of freedom in fitting to actual conditions by adding or subtracting a constant amount to each loss before taking the logarithm and fitting the curve in the manner described. Arthur L. Bailey uses this distribution in much of his work. Robert A. Bailey[17] has found indications that the log-normal distribution is approximate also to fire insurance.

In our second model we assume that the probability of accident in any

[17] Experience Rating Reassessed—CAS XLVIII, (1961)

period of time is the same for each individual exposure and that it is proportional to the time exposed as in our first model; but instead of assuming that the amount of each loss is the same, we assume that the amount of each loss is distributed by some frequency curve which may be expressed as a log-normal curve or some other appropriate curve. The effect of this assumption is to increase considerably the number of claims required for full credibility of pure premiums and other functions involved in rate making. The mathematical development of these credibility standards is set out in Appendix C. As already explained, this does not alter the standards of full credibility used in practice, but it is important, for example, in showing that a trend factor based on two averages of equal volume requires twice as many claims in each average to produce the same level of credibility.

With models of this nature, we can approach more complex problems of relative credibilities such as the credibility requirements for various types of fire insurance classifications. Robert L. Hurley[10] develops a table of credibilities for Dwellings, Mercantile Contents and Manufacturing fire insurance risks which specifies the following volumes of premium for full credibility:

Dwellings	$ 2,000,000
Mercantile Contents	16,000,000
Manufacturing	60,000,000

The importance of these figures lies not in their absolute amounts, but in showing how much more readily dwelling experience acquires full credibilty than does the experience of commercial risks.

The development of the basic formulae for the distribution of insurance statistics due to chance fluctuations only, when we have a skew distribution of losses by size combined with a distribution of the number of losses according to the Poisson formula, has been worked out in detail by Arthur L. Bailey[12].

11. CREDIBILITY AND EXPERIENCE RATING PLANS

The problem of experience rating arises out of the necessity, from the standpoint of equity to the individual risk, of striking a balance between class-experience on the one hand and risk experience on the other. —Albert W. Whitney[9]

Experience rating plans, which first developed in the rating of Workmen's Compensation Risks, are as old as the Casualty Actuarial Society itself and are an important application of actuarial theory to insurance rate making. Such plans are used in practically all branches of casualty insurance and have recently been used for various types of property coverage.

Arthur L. Bailey[12] has pointed out that there are two kinds of credibility; the one we have so far discussed for rate revisions and the one used in experience rating plans. He calls the first the "limited fluctuation credibility" and the other the "greatest accuracy credibility". While the purpose of the formulae used in these two applications of credibility are not the same, it is difficult to accept such a simple definition in either case. While limiting fluctuations is important in rate revisions, responsiveness to trends is even more important and higher standards for full credibility would be used if we were concerned only with rate stability. Experience rating plans are so varied that one cannot help concluding that competitive expediency has played an important part in their design and there is some doubt that the majority of formulae give sufficient credibility to the individual experience to justify the appellation "most accurate".

Experience rating plans provide an adjustment in manual rates to reflect the experience of the individual risk; in the most usual form of plan the premium charged is in the form

$$\text{Manual Premium x}\left[1 - Z + Z\,\frac{A}{E}\right]$$

where A and E are the actual and expected losses and Z is the credibility factor. It will be noted that when $Z = 0$ the manual premium is charged; when $Z = 1$ the manual premium is adjusted in the ratio of the actual experience of the risk to expected experience of the risk under the manual premium plan. In the latter case the risk is said to be self-rated. The formulae actually used in experience rating plans are not so simple as this because they normally contain adjustments to reduce the effect of individual large losses.

Arthur L. Bailey[2] describes these adjustments as follows:

> "In addition to the relatively simple concept that more consideration or weight should be given to a greater volume of observational data, the casualty actuaries have devised credibility procedures to give more weight to the frequent occurrence of small losses than to the occasional or fortuitous occurrence of large losses of the same total amount. (It should be noted that negative losses cannot occur.) For example, the rate making procedure for workmen's compensation insurance separates the actual losses into 'Serious', 'Non-Serious' and 'Medical' losses and uses three differing schedules of credibility for the three components of the total loss. Several experience rating plans give a greater schedule of credibility to the first G dollars of each loss than is given to the excess of any loss over G dollars. The 'Multi-Split Experience Rating Plan' for workmen's compensation insurance carries this even further by providing, in effect, a separate schedule of credibilities for each interval of G dollars of which a loss is composed".

A discussion of the details of various experience rating plans would be out of place in this outline, but Robert A. Bailey[17] has set down the following criteria for such plans:

I. 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.

   A number of other criteria are imposed which are in the nature of limitations on this fundamental criterion. They are:

II. The risk's premium should not fluctuate widely from year to year. If it fluctuates too widely, the purpose of insurance is defeated.

III. One dollar of actual loss should not increase the adjusted losses by more than one dollar.. Otherwise the insured might find it to his advantage to pay his own losses. (The term "adjusted losses" means the weighted average of the actual and expected losses which is used to determine the adjusted rate for the risk.)

IV. The experience rating plan should not be too expensive to administer.

It is also desirable to quote the conditions which the credibility Z should satisfy as formulated by Francis S. Perryman[18].

(i) The credibility should be not less than zero and not greater than unity.

(ii) The credibility should increase (or more strictly speaking not decrease) as the size of the risk increases.

(iii) As the size of the risk increases the percentage charge for any loss of given size should decrease.

The Educational Committee of the Casualty Actuarial Society is preparing a students' guide to Experience Rating which will provide a fuller introduction to this important field.

12. VARIATIONS IN INHERENT LOSS FREQUENCY—THE THIRD MODEL

It is recognized that individual risks within a classification are not alike and that there exist inherent differences . . . These differences are of such a nature that it is difficult to label them definitely and they cannot be associated with conditions measureable in advance.

—Paul Dorweiler[19]

In order to establish rating plans data are classified into a large number of breakdowns. For example in private passenger automobile insurance, there are classifications by state, territory within the state, type of automobile, use of automobile, age of automobile and age and sex of driver. Rates are established for each combination of these classifications; and in the models we have so far developed, we assume that the probability of accident is the same for all exposures in any single combination of classifications. The probability of accident (all other factors being equal) will not vary by a marked jump as we proceed over the boundary line from one territory to another but will vary continuously as we move across the state. Practical necessity calls for the use of a limited number of classifications which are chosen on a judgment basis to provide groups of reasonable homogeneity, but it is clear that there must be variation in the true probability within a single classification group. However, there is every reason to believe that there is considerably more heterogeneity in each group than that suggested by the above argument. The criteria used to determine classifications are not the only possible ones. For example mileage, horsepower, occupation and many other classifications are possible in automobile insurance. There can be no question that in most cases an actual classification will embrace quite a wide distribution of probabilities of accident. For this reason we assume in our third model that the probability of accident within a classification is not fixed but is distributed over a range defined by some frequency curve.

It is usual to assume that the distribution of probabilities of accident within the classification follow a Pearson Type III curve because this is a

[18] Experience Rating Credibilities—CAS XXIV, p. 60 (1937)

[19] Presidential Address—CAS XXI, p. 1 (1934)

skew form and because it leads to a conveniently simple equation for fitting. It is further assumed that the probability for a given individual remains constant throughout the experience period. The result of this is to replace the Poisson Distribution by the Negative Binomial Distribution in the model.[20, 21, 22] It is sometimes possible to experiment with this more accurate model and to avoid, at least in preliminary studies, the rather extensive arithmetic of the negative binomial distribution, by substituting a three-point or five-point probability distribution.[23] Thus we may assume ½ the exposures have a probability of accident within one year of .10; ¼ a probability of .05 and ¼ a probability of .15. A skew distribution can, of course, be used. While this method is often helpful in preliminary studies, it may, when used by the inexperienced, suggest misleading results.

In testing certain actual automobile experience against the model, using the negative binomial distribution, excellent agreement between actual and theoretical distributions was observed.[24] This model has been found to be particularly helpful in the field of merit rating discussed in the next section.

13. MERIT RATING

In writing private passenger automobile liability insurance there has always been a need for underwriters to select the good business and turn down the poor because the rate classification system has never been perfect.

—Robert A. Bailey[25]

In the third model described in the preceding section, we have assumed that within each classification there is quite a wide range of variation in the probability of loss; we have also assumed that the probability remains constant for a given individual. This suggests a new form of classification which depends on the loss history of the individual and varies the classification according to the period elapsed since the last loss or to the number of losses in a recent period of time. It is not difficult to show that, on the basis of the model, significantly different class rates will develop for risks classified in this manner. Variation of rates according to loss history is used in private passenger automobile, homeowners and other lines of insurance and is called merit rating. Such rating is usually associated with other forms of classifications. Sometimes merit rating is determined not only by actual loss history, but also by a combination of loss history and of some data closely correlated to the potential loss experience such as traffic violation records. It is necessary, of course, to test the appropriateness of the model

[20] Lester B. Dropkin, Some Considerations on Automobile Rating Systems Utilizing Individual Driving Records—CAS XLVI, p. 165 (1959)

[21] LeRoy J. Simon, The Negative Binomial and Poisson Distributions Compared—CAS XLVII, p. 20 (1960)

[22] Charles C. Hewitt, Jr., The Negative Binomial Applied to the Canadian Merit Rating Plan for Individual Automobile Risks—CAS XLVII, p. 55 (1960)

[23] This approach is similar to the n-ages method of approximate valuation in life insurance.

[24] LeRoy J. Simon, Fitting Negative Binomial Distributions by the Method of Maximum Likelihood—CAS XLVIII, (1961)

[25] Any Room Left for Skimming the Cream?—CAS XLVII, p. 30 (1960)

against the actual development of experience under a merit rating plan and this test has proved satisfactory.

Considerable misunderstanding exists about the principles of merit rating[26] because of failure to realize that merit rating is a system of classification to which the normal credibility criterions for rate making apply. A merit rated risk is one of a large class of similar risks all of which meet certain classification standards including one defined in terms of past loss experience. Normal rate making methods can be used to develop the correct rate relativities under such a plan. Such relativities are called merit credits and debits.

Robert A. Bailey[25] in the paper from which the quotation at the head of this section is taken discusses some of the impact of merit rating upon existing classification plans. A point not touched upon is that since the distribution of loss frequency within the separate territorial and other classifications is not uniform, the introduction of merit rating, particularly if the credits and debits are large, may well lead to a reduction in the rate differentials required for territorial and other classifications and could possibly lead to a simplification in the overall classification system.

A number of important papers on merit rating have been published in recent volumes of the Proceedings of the Casualty Actuarial Society; and reference should be made to these for further discussion of this aspect of credibility.[16, 20, 27, 28, 29, 30]

The following conclusions from one of these studies[28] provides a fitting ending to this section.

"In summary, we feel that the Canadian merit rating data for private passenger cars leads to the following conclusions:

(1) The experience for one car for one year has significant and measurable credibility for experience rating.

(2) In a highly refined private passenger rating classification system which reflects inherent hazard, there would not be much accuracy in an individual risk merit rating plan, but where a wide range of hazard is encompassed within a classification, credibility is much larger.

(3) If we are given one year's experience and add a second year we increase the credibility roughly two-fifths. Given two years' experience, a third year will increase the credibility by one-sixth of its two-year value."

---

[26] LeRoy J. Simon, Merit Rating Myths and Mysteries—Automobile Insurance Rate Making. Casualty Actuarial Society, 1961

[27] Herbert E. Wittick, The Canadian Merit Rating Plan for Individual Automobile Risks —CAS XLV, p. 214 (1958)

[28] Robert A. Bailey and LeRoy J. Simon, An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car—CAS XLVI, p. 159 (1959)

[29] Frank Harwayne, Merit Rating in Private Passenger Automobile Liability Insurance and the California Driver Record Study—CAS XLVI, p. 189 (1959)

[30] Lester B. Dropkin, Automobile Merit Rating and Inverse Probabilities—CAS XLVII, p. 37 (1960)

14. REINSURANCE, SURPLUS PROBLEMS, ETC.

The object of the theory of risk is to give a mathematical analysis of the random fluctuations in an insurance business and to discuss the various means of protection against their inconvenient effects.

—Harold Cramer[31]

The third model introduced earlier provided for variations of losses by size and for variations in the individual probabilities of loss. We have so far applied this model to discuss classification rate making, but the same or similar models can be applied to a whole class of business or, indeed, to the whole portfolio of an insurance company. Such a model can be used to study many problems in reinsurance, particularly excess of loss and stop loss coverages. Again it can be used to determine retention limits when these are dependent on capacity alone and not on underwriting or other considerations.

The model, with appropriate developments, can be further used to discuss surplus requirements and similar problems. There is an extensive literature on this subject under the general title of the Theory of Risk, mostly published in Western Europe. No attempt has been made to include these writings in the bibliography in Appendix A.

15. CONCLUSION

—the business finds itself with still a large number of problems on its hands, many of which we know the actuary will eventually have to solve. Let him, therefore—the casualty actuary about whom I have been talking—continue to grapple with these problems, knowing full well that he has an enormous advantage in the possession of a scientific mind and of scientific methods; with these he will, on his merits, be called on to play a larger and most responsible part in the business of casualty insurance.

—Francis S. Perryman[32]

The above remarks from a Presidential Address to the Casualty Actuarial Society in 1939 are as true today as when they were spoken, further the duties of the actuary now extend to the property insurance as well as to casualty insurance. In this brief outline, an attempt has been made to provide the reader with a simple framework into which he can place the very large number of important contributions to credibility theory which have appeared in the Proceedings of the Casualty Actuarial Society. There is still much opportunity for original and important research in the field, even in the areas covered by the models already discussed. Particularly in the area of experience rating, much work and testing is needed.

Except for its application to merit rating, little practical use has been made so far of the third model, which recognizes the diversity of risks within an individual classification. Considerable development of the mathematics required for this field of study, when the skew distribution of losses by size

[31] On the Mathematical Theory of Risk, Skandia Jubilee Volume, Stockholm (1930)
[32] Presidential Address—CAS XXV, p. 291 (1939)

is also considered, has been undertaken by Arthur L. Bailey[12] and further study should lead to important developments.

The third model assumes that the loss frequency distribution is not correlated in any way with the probability of loss. It seems probable that within any classification group, the average amount of an individual loss of those persons with a low loss frequency may well be lower than the general average amount of loss for the group as a whole. The study and testing of models reflecting this idea may well lead to larger merit rating credits than are at present customary in the United States and Canada. Again, it is not entirely true that the probability of an accident remains constant for any one individual. The fluctuation which occurs in use of an automobile over the year must be reflected in the probability of accident. Further, an automobile driver is likely to show special caution in a short period immediately following an accident and only slowly return to his pre-accident standards. Also, since there is a correlation between age of driver and accident proneness, there must be trends in probability of an individual having an accident. The construction *and testing* of new models is an important field of actuarial study, which will provide one of the most powerful means of attacking those problems which are still unsolved and the new problems which will arise in the future.

In conclusion the development of credibility theory is one of the more important aspects of actuarial science. Much has been accomplished in the nearly fifty years since the formation of the Casualty Actuarial Society, but more remains to be done. It is hoped that this brief outline will help the reader to obtain a grasp of the principles involved, or, if he is already familiar with the subject, to reassess some of the problems still awaiting final solution. It is perhaps necessary to stress that credibility procedures are not a substitute for informed judgment, but an aid thereto. Of necessity so many practical considerations must enter into any actuarial work that the student cannot substitute the blind application of a credibility formula for the careful consideration of all aspects of an actuarial problem.

Appendix A

Bibliography

I. *Early papers which foreshadow the later development of Credibility Theory*

FARREN, E. J., On the Improvement of Life Contingency Calculation. Journal of the Institute of Actuaries, Vol. 5, p. 185 (1855)

KANNER, DR. M., On the determination of the Average Risk attaching to the grant of Insurance upon Lives. Journal of the Institute of Actuaries, Vol. 14, p. 439 (1869)

WOOLHOUSE, W. S. B., On the Philosophy of Statistics. Journal of the Institute of Actuaries, Vol. 17 (1873)

WILLETT, A. H., The Economic Theory of Risk and Insurance. 1901 (Reprinted Philadelphia 1951)

VARIOUS The Problem of Mathematical Risk; special reserves of Insurance Companies and Pension Funds. VI International Congress of Actuaries. Vienna. 1909

II. *More important papers on Credibility in the Proceedings of the Casualty Actuarial Society*

MOWBRAY, A. H., How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium? CAS I, p. 24 (1914)

WOODWARD, J. H., The Experience Rating of Workmen's Compensation Risks. CAS II, p. 356 (1916)

FISHER, ARNE, Outline of a Method for Determining Basic Pure Premiums. CAS II, p. 394 (1916)

MOWBRAY, A. H., A New Criterion of Adequacy of Exposure. CAS IV, p. 263 (1918)

WHITNEY, A. W., The Theory of Experience Rating. CAS IV, p. 274 (1918)

MICHELBACHER, G. F., The Practice of Experience Rating. CAS IV, p. 293 (1918)

WHEELER, R. A., Credibility and Automobile Rate Making. CAS XVI, p. 268 (1930) (Reprinted in Automobile Insurance Rate Making)

PERRYMAN, F. S., Some Notes on Credibility. CAS XIX, p. 65 (1932) (Reprinted in Automobile Insurance Rate Making)

DORWEILER, PAUL, A Survey of Risk Credibility in Experience Rating. CAS XXI, p. 1 (1934)

PERRYMAN, F. S., Experience Rating Plan Credibilities. CAS XXIV, p. 60 (1937)

SMICK, J. J., Merit Rating—The Proposed Multi-Split Experience Rating Plan and the Present Experience Rating Plan. CAS XXVI, p. 84 (1939)

JOHNSON, R. A., JR., The Multi-Split Experience Rating Plan in New York. CAS XXVIII, p. 15 (1941)

BAILEY, A. L., Sampling Theory in Casualty Insurance. CAS XXIX, p. 50 and CAS XXX, p. 31 (1942-3)

SATTERTHWAITE, F. E., Notes on Mathematical Statistics. CAS XXIX, p. 122 (1942)

BAILEY, A. L.,	A Generalized Theory of Credibility. CAS XXXII, p. 13 (1945)
CARLETON, JOHN,	Non-Random Accident Distributions and the Poisson Series. CAS XXXII, p. 21 (1945)
SCAMMON, L. W.,	Massachusetts Workmen's Compensation Rate Making—Primary—Excess Basis. CAS XXXIV, p. 17 (1947)
GROSSMAN, E. A., and FRIEDMAN, BERNARD	Accident Rates with Confidence Limits. CAS XXXIV, p. 95 (1947)
BAILEY, A. L.,	Workmen's Compensation D-Ratio Revisions. CAS XXXV, p. 26 (1948)
BAILEY, A. L.,	Credibility Procedures. CAS XXXVII, p. 7 (1950) (See also discussion CAS XXXVII, p. 94)
KORMES, MARK,	Notes on Experience Rating Credibility. CAS XXXIX, p. 98 (1952)
HURLEY, R. L.,	A Credibility Framework for Gauging Fire Classification Experience. CAS XLI, p. 161 (1954) (See also discussion CAS XLII, p. 241) (Reprinted in Fire Insurance Rate Making and Kindred Problems)
WITTICK, H. E.,	The Canadian Merit Rating Plan for Individual Automobile Risks. CAS XLV, p. 214 (1958) (Reprinted in Automobile Insurance Rate Making)
BAILEY, R. A., and SIMON, LER. J.,	An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car. CAS XLVI, p. 159 (1959) (Reprinted in Automobile Insurance Rate Making)
DROPKIN, L. B.,	Some Considerations on Automobile Rating Systems Utilizing Individual Driving Records. CAS XLVI, p. 165 (1959) (Reprinted in Automobile Insurance Rate Making)
HARWAYNE, FRANK,	Merit Rating in Private Passenger Automobile Liability Insurance and the California Driver Record Study. CAS XLVI, p. 189 (1959) (Reprinted in Automobile Insurance Rate Making)
HURLEY, R. L.,	Multiple Peril Rating Problems. CAS XLVI, p. 196 (1959)
ROBERTS, L. H.,	Credibility of 10/20 Experience as Compared with 5/10 Experience. CAS XLVI, p. 235 (1959) (Reprinted in Automobile Insurance Rate Making)

UHTHOFF, D. R.,	The Compensation Experience Rating Plan. CAS XLVI, p. 285 (1959)
BAILEY, R. A. and SIMON, LER. J.,	Two Studies in Automobile Insurance Ratemaking. CAS XLVII, p. 1 (1960)
SIMON, LER. J.,	The Negative Binomial and Poisson Distributions Compared. CAS XLVII, p. 20 (1960)
BAILEY, R. A.,	Any Room Left for Skimming the Cream? CAS XLVII, p. 30 (1960) (Reprinted in Automobile Insurance Rate Making)
DROPKIN, L. B.,	Automobile Merit Rating and Inverse Probabilities. CAS XLVII, p. 37 (1960) (Reprinted in Automobile Insurance Rate Making)
HEWITT, C. C., JR.,	The Negative Binomial Applied to the Canadian Merit Rating Plan for Individual Automobile Risks. CAS XLVII, p. 55 (1960)
KLAASSEN, E. J.,	Multiple Coverage Experience Rating Plan. CAS XLVII, p. 66 (1960)
SIMON, LER, J.,	Fitting Negative Binomial Distributions by the Method of Maximum Likelihood. CAS XLVIII, (1961)
BAILEY, R. A.,	Experience Rating Reassessed. CAS XLVIII, (1961)

III. *Some other papers related to Credibility*

KEFFER, RALPH,	An Experience Rating Formula. Transactions of the Actuarial Society of America. Vol. 30, p. 130 (1929)
LONGLEY-COOK, L. H. and PRUITT, D. M.,	Law of Large Numbers in the Fire Insurance Business and Credibility of Statistics Used. Proceedings Insurance Accounts Association. (1952) (Reprinted in Summary in Snider's Readings in Property and Casualty Insurance)
BORCH, K.,	An Attempt to Determine the Optimum Amount of Stop Loss Reinsurance. XVIth International Congress of Actuaries I, p. 597 (1960)
AMMETER, H.,	Stop Loss Cover and Experience Rating. XVIth International Congress of Actuaries I, p. 649 (1960)
BEARD, R. E.,	Analytical Expressions of the Risks Involved in General Insurance. XVth International Congress of Actuaries II, p. 230 (1957)

KELLY, A. B., How Factory Mutual Rates are Established. National Insurance Buyer (Nov. 1957) (Reprinted in Snider's Readings in Property and Casualty Insurance)

SIMON, LER. J., Merit Rating Myths and Mysteries—Automobile Insurance Rate Making.

IV. *Mathematical Theory of Risk*

A bibliography of the Mathematical Theory of Risk is being prepared by a Committee of the Casualty Actuarial Society.

Appendix B

The Poisson Distribution[33]

If the time interval of exposure is made sufficiently small the number of claims (including multiple claims as single claims) arising from a single exposure unit will be either zero or one. The distribution of the sum of the claims from all exposure units is then described by the binomial distribution.

Let q represent the true probability of a claim occurring in one year, and let n represent the number of exposure units with the yearly claim frequency q.

Then, the probability that the total number of claims will be exactly r out of ns trials, where s is large enough so that for each of the n exposure units not more than one claim will occur in time interval 1/s, is given by the (r + 1) th term in the expansion of the binomial $[(1 - q/s) + q/s]^{ns}$, where q/s is the probability of the occurrence of one claim in the time interval 1/s. This (r + 1) th term is

$$ {}_{ns}C_r \, (1 - q/s)^{ns-r} \, (q/s)^r $$

The probability P that the number of accidents in ns trials will be within ± 100k% of nq (= ns.q/s = the expected value of r) is therefore equal to

$$ \sum_{r=(1-k)nq}^{r=(1+k)nq} {}_{ns}C_r \, (1 - q/s)^{ns-r} \, (q/s)^r $$

By using Stirling's formula for factorials, it can be shown that this expression is approximated by

$$ P = \frac{2}{\sqrt{2\pi\, nq}} \int_0^{knq} e^{\frac{-x^2}{2nq}} \, dx $$

as s becomes very large, and where n is large and k is not large.

[33] From a memorandum prepared by the National Bureau of Casualty Underwriters in 1949.

It may be shown that as s becomes large

$$ {}_{ns}C_r \, (1 - q/s)^{ns-r} \, (q/s)^r \text{ approaches } \frac{e^{-nq}\,(nq)^r}{r!} $$

which is the general term of the Poisson distribution of r with expected value nq. Thus, the expression for P is equally valid under the assumption that the number of claims has a Poisson distribution.

Change the variable to $t = x/\sqrt{nq}$, so that

$$ P = 2\left[\frac{1}{\sqrt{2\pi}}\int_0^{k\sqrt{nq}} e^{-\frac{t^2}{2}}\,dt\right] $$

The upper limit of the integral in the brackets corresponding to a given value of P may be read from tables of the standard normal integral. Then the expected number of claims, nq, may be calculated for any given value of k. For example, if P = .95, a table of values of the integral shows that

$$ k\sqrt{nq} = 1.960, \text{ and if } k = .075 $$

$$ nq = \left(\frac{1.960}{.075}\right)^2 = 683 \text{ claims.} $$

Appendix C

Relationship of Credibility Standards for Claim Frequencies, Claim Costs, Pure Premiums and Trends[34]

The credibility tables commonly used in rate making are developed as the credibility of the claim frequency. For example, if the expected number of claims is 1082, the actual number will be within 5% of 1082 90% of the time. The formula is based on the Poisson distribution and the number of claims for 100% credibility is derived from the formula P^2/K^2 where $\pm$ P are the values of the normal curve corresponding to a selected probability, and K is the selected deviation from the expected. In the example cited above, the probability is 90%, P = 1.645 and K = .05.

If we need P^2/K^2 claims to provide a selected level of credibility for the claim frequency, how many claims do we need to provide the same level of credibility for the claim cost, the pure premium and trends?

DEFINITION OF SYMBOLS

N	= number of intervals each exposure year is divided into
Y	= number of exposure years
NY	= number of exposure intervals
M_f	= average claim frequency per exposure interval
M_c	= average cost per claim
M_{pp}	$= M_f M_c =$ average cost per exposure interval

[34] From a memorandum prepared by Robert A. Bailey

S_f^2 = variance of claim frequency
S_c^2 = variance of claim cost
S_{pp}^2 = variance of pure premium per exposure interval
NY M_f = expected number of claims
P = value on the normal curve corresponding to a selected probability
K = selected deviation from the expected average.

The number of claims which provides the selected level of credibility is determined as follows:

Claim Frequency

$$KM_f + S_f/\sqrt{NY} = P$$
$$NYM_f = P^2/K^2 \times S_f^2/M_f$$
$$NYM_f = P^2/K^2, \text{ since } S_f^2 = M_f$$

Claim Cost

$$KM_c + S_c/\sqrt{NYM_f} = P$$
$$NYM_f = P^2/K^2 \times S_c^2/M_c^2$$

Pure Premium

$$KM_{pp} + S_{pp}/\sqrt{NY} = P$$
$$NYM_f = P^2/K^2 \times S_{pp}^2/M_c^2 M_f$$
$$NYM_f = P^2/K^2 \times (1 + S_c^2/M_c^2)$$

since

$$S_{pp}^2 = \Sigma C^2/NY - (\Sigma C/NY)^2$$

but $(\Sigma C/NY)^2$ becomes insignificant as N is increased therefore

$$S_{pp}^2 = \Sigma C^2/NY$$
$$S_{pp}^2 = M_f\left[(\Sigma C/NYM_f)^2 + \Sigma C^2/NYM_f - (\Sigma C/NYM_f)^2\right]$$
$$S_{pp}^2 = M_f(M_c^2 + S_c^2)$$

Trends

The trend is actually a measurement of the difference between two averages. The variance of the difference between two independent variables is the sum of the variances of the two variables. This means that if two averages are based on approximately the same volume of experience, twice as many claims are needed in each average to produce the same level of credibility in the difference between them. If one average is based on twice or four times as much experience, then 1½ or 1¼ as many claims, respectively, are needed in the average based on the fewer claims to produce the same level of credibility in the difference.

Summary

If the expected number of claims needed to produce a selected level of credibility for the claim frequency is taken as 1, the number required to produce the same level of credibility is S_c^2/M_c^2 for the claim cost, and 1 plus S_c^2/M_c^2 for the pure premium. S_c^2/M_c^2 can be determined from a distribution of claims by size of claim and for most lines of insurance it ranges between 2 and 4. A trend factor based on two averages of equal volume requires twice as many claims in each average to produce the same level of credibility.