DISCUSSION OF THE RATEMAKING PROCEDURE IN WORKMEN'S COMPENSATION INSURANCE

A Method of Testing Classification Relativities

BY

STEFAN PETERS

A. *Introduction*

If in one of the natural sciences a scientist is to study a complex phenomenon which is determined by many elementary causes he usually approaches his problem in three different ways. He first studies the phenomenon in a purely empirical manner trying to describe his measurements by means of a mathematical formula or graph. He then tries to develop a theory regarding the action of the many small causes which in the aggregate produce the phenomenon under investigation, and finally he tests independently his hypothesis regarding the elementary causes. Thus, in thermodynamics the specific heats of gases, as determined by experiments, are first described as a function of the molecular composition of these gases, then a hypothesis is developed relative to the action of the individual molecules (e.g. the kinetic theory of gases) and finally the assumptions regarding the behavior of the individual molecules are tested by the physicist through independent experiments.

The problems facing the research worker in natural sciences are not essentially different from those facing the actuary in casualty insurance when concerned with the task of making rates for a great number of different classifications. The method followed by the actuary, however, is not so complete as that followed by, say, the physicist. He describes the composite phenomenon, for instance, he determines the required rate level from experience; he develops a theory as to the individual causes producing the general phenomenon, that is, he estimates classification relativities, credibilities and law amendment factors on the basis of certain statistical assumptions; but he usually fails to test his hypothesis independently as does the physicist. This paper is intended to complete the actuarial method by testing the various assumptions made in the computation of rates. The subject is restricted to the ratemaking procedure in workmen's compensation insurance, but it is believed that some of the methods proposed will be applicable —with due changes—to other casualty lines.

The present part of the paper, in particular, is concerned with a method of testing classification relativities. No attempt has been made to draw definite conclusions as to the accuracy and usefulness of the present procedure of selecting pure premiums. This may perhaps be done in a later part of this paper to be published in the future. The main purpose of this part of the paper is to develop a *method* of testing a given set of selected pure premiums as to its accuracy, or better, of comparing two different sets of selected pure premiums as to their relative accuracy. Since the approach to this problem is new, and this study, therefore, cannot benefit from past experience, the method proposed will doubtlessly contain many faults and be subject to improvement, suggestions for which, the author hopes, will be forthcoming in the discussion of this paper.

B. *The Present Method of Determining Classification Relativities*

Under the present ratemaking program, classification relativities are determined by the computation, for each state, of a set of selected pure premiums at the time of a proposed rate revision. For classifications with a large volume of exposure these pure premiums are simply equal to the indicated pure premiums obtained from the actual state experience incurred under these classifications during the last five policy years after the losses for different policy years have been brought to a common level. These pure premiums are determined separately for serious, nonserious and medical losses. For classifications with a small volume of exposure for which the limited volume of experience does not permit reliance entirely upon the indications of the actual experience for a five year period, these indicated premiums are weighted against national pure premiums. Thus a set of formula pure premiums is obtained which generally is selected as the final pure premium for the classification. Occasionally, however, the formula pure premium is modified either by judgment or according to certain rules which it is not necessary to mention here. The weight accorded to the state indications of the classifications increases with increasing size of expected losses. 100% credibility is assigned to the state indications for serious pure premiums where the expected losses are equal to or larger than 25 times the average death and permanent total indemnity loss; for non-serious pure premiums the criterion for 100% credibility is expected losses of at least 300 times the average cost of a non-serious case; and for medical pure premiums the criterion for 100% credibility is expected losses of at least 80% of the non-serious criterion for 100% credibility. Between 100% and 0% the credibility is considered a linear function of the expected losses but only credibility values of 100%, 75%, 50%, 25%, 15%, 10% and 0% are used.

The national pure premiums forming a part of the selected pure premiums for classifications with small exposure are derived from national pure premiums on a basic level. These are computed from countrywide experience for five policy years, brought to the common basic level by application of conversion factors which are determined separately for serious, non-serious and medical losses and for each policy year. In ratemaking, the national pure premiums on the basic level are brought back to the experience level of the state for which they are to be used by means of reversion factors which are calculated separately for serious, non-serious and medical pure premiums and for each industry group. These reversion factors are determined in such a way that, separately by industry group and by parts, the aggregate expected losses derived from national pure premiums total to the same amount as that portion of the expected losses derived from indicated state pure premiums which they replace in the formula expected losses.

The author does not intend to give a detailed critical analysis of the theory underlying this method of computing selected pure premiums. Some of the more important objections which have been raised against the method will briefly be mentioned. In many instances it is doubtful whether the experience incurred under the same classification in different states can properly be combined since the nature of the operations covered under the same classification frequently differs to a substantial degree. Another objection is caused by the difference in the nature of the conversion factors used in assembling the experience required for the computation of national pure premiums on the basic level and of the reversion factors used in reverting the national pure premiums to any particular state level. This has the effect of distorting the pure premium for those classifications, the bulk of whose experience comes from one state only, if the industry group rate level differs substantially from the rate level for all industry groups combined. Another point raised is that the reversion factor from the national to state level depends to a much higher degree on the experience of classifications with a small volume of exposure than on that of classifications with a large volume of exposure, as can be seen from the detailed formulas, and thus has been based on a relatively small loss volume which also is subject to large casual fluctuation.

These and several other reasons make it appear probable that the present ratemaking procedure can be substantially improved and it is the purpose of this paper to furnish the tools which enable the actuary to decide whether any given set of pure premiums is better (or worse) than the set of selected pure premiums determined by the present ratemaking procedure.

C. *Theory of the Proposed Method of Analysis*

If a set of pure premiums is to be tested for its accuracy, the obvious approach is to compare the expected losses produced by these pure premiums with the actual losses for a sufficiently long period of time so that one may expect the actual losses to be only slightly influenced by chance fluctuations and to present a close estimate of the "true" expected losses. This course, unfortunately, is difficult to follow in testing the pure premiums for workmen's compensation insurance, firstly, because the experience required for the classifications with small exposure would have to extend over a very long period of time and may even not be available, and secondly, because the combination of the experience for widely separated policy years presents peculiar difficulties due to changes in the benefit level and in the scope of classifications which would make it necessary to use certain assumptions in order to be able to combine the experience incurred in different periods.

The large number of classifications, however, permits another approach to the problem. If we use only the experience of one policy year, but look at the errors due to the method of selecting pure premiums as fortuitous events which are independent for different classifications, we can consider some appropriate quantities which measure these errors as the elements of frequency distributions and base our test on the analysis of these frequency distributions. Details of the manner in which this is accomplished are given later. Since the "true" pure premiums for serious, non-serious and medical losses of the various classifications are not known, we cannot actually measure the error due to the method of selecting a given pure premium, but we can only measure the deviation of the expected losses based on the selected pure premiums from the actual losses incurred during the policy year under consideration. Our frequency distribution of these deviations will, therefore, measure the composite effect of (1) the deviation of the expected losses based on the selected pure premiums from the "true" expected losses which is due to the method of selecting pure premiums and of (2) the deviation of the actual losses from the "true" expected losses which is entirely due to chance.

The quantity which presents itself to mind at first consideration as the most convenient measure of the deviation of actual from expected losses is the ratio of actual to expected losses. This ratio is evidently a positive number which varies from zero to very large amounts and the frequency distribution of this ratio, whose mean must be in the neighborhood of unity, would, therefore, necessarily be skewed. In order to obtain a symmetrical frequency distribution which would be easier to work with, the logarithm (with a base 10) of this ratio has been chosen as a measure of the deviation of actual from expected losses and the frequency distributions thus obtained are actually symmetrical for all practical purposes.

The measure of the deviation of actual from expected losses which will be used in this study, is therefore

$$x = \log_{10} \frac{\text{actual losses}}{\text{expected losses}} \tag{1}$$

This can also be written

$$x = x_a + x_n \tag{2}$$

where

$$x_a = \log_{10} \frac{\text{actual losses}}{\text{“true” expected losses}} \tag{3}$$

$$(4) \qquad x_n = -\log_{10} \frac{\text{expected losses based on selected p.p.'s}}{\text{"true" expected losses}}$$

As the "true" expected losses are not known, the resolution of x into its component parts x_a and x_n is only theoretical and cannot actually be achieved. Although our interest is concentrated exclusively on x_n, we can only study the distribution of $x = x_a + x_n$.

The unknown distributions of x_a and x_n are evidently independent of each other since the one depends on the chance fluctuations of actual losses and the other on the method of selecting pure premiums. The variance σ^2 of the quantity $x = x_a + x_n$ will therefore be equal to the sum of the variance σ_a^2 of x_a and the variance σ_n^2 of x_n.

$$(5) \qquad \sigma^2 = \sigma_a^2 + \sigma_n^2$$

The better the selected pure premiums fit the "true" pure premiums the smaller the variance σ_n^2 will be. The variance σ_a^2 of x_a, however, does not depend on the method of selection of pure premiums. If we, therefore, compare the distributions of x_1 and x_2 for two different sets of selected pure premiums, the variances σ_1^2 and σ_2^2 of x_1 and x_2 will be composed of one common item σ_a^2 corresponding to x_a and an additional item $\sigma_{n,1}^2$ or $\sigma_{n,2}^2$ respectively corresponding to $x_{n,1}$ and $x_{n,2}$:

$$(6) \qquad \sigma_1^2 = \sigma_a^2 + \sigma_{n,1}^2 \qquad\qquad \sigma_2^2 = \sigma_a^2 + \sigma_{n,2}^2$$

Since that variance $\sigma_{n,i}^2$ which corresponds to the pure premiums with the better fit will be smaller than the other, the same will be true for the corresponding σ_i^2. It is on this principle that our method of testing two sets of selected pure premiums against each other is based. The set producing the smaller variance will—if the reduction in the variance is large enough not to be attributable to chance—be the set which comes nearest to the ideal, yet not determinable, "true" pure premiums.

It is possible for one or both of the sets of selected pure premiums to contain a systematic bias. This would show up in the mean of the frequency distribution of x, however, and an analysis of the means will, therefore, be made preceding the analysis of the variances.

As the aim of this study is to test the accuracy of the selection of pure premiums for all classifications, irrespective of whether the effect of any deviation of the selected from the "true" pure premiums relative to the total premium volume is large or small, the indications for classifications with a large volume of exposure were not assigned greater weight than those for classifications with a small volume of exposure. Since, however, that part of the deviation of actual losses from expected losses which is caused by the chance fluctuation of actual losses will obviously be distributed with larger dispersion for classifications with small exposure than for classifications with a larger exposure, the frequency distributions of x have been determined separately for classifications whose state credibility under the present ratemaking procedure is 50% or over and for classifications whose state credibility is less than 50%.

D. *The Computation of the Logarithmic Deviations*

The test method outlined above has been applied to the selected pure premiums for the July 1, 1938 rate revision in New York through the use of the expected losses obtained by extending policy year 1938 statutory medical coverage payrolls at selected medical pure premiums and policy year 1938 total payrolls at selected serious indemnity and non-serious indemnity pure premiums. These expected losses were compared with the actual losses incurred during policy year 1938 in New York, as shown in the exhibits of classification experience which were prepared for the July 1, 1941 rate revision.* Vessel classifications, special New York classifications, "*a*" rated classifications which are rated according to the nature of the operations and chemical classifications were excluded from the material used since they are subject to a special ratemaking procedure.

The policy year 1938 actual losses were chosen for a comparison with the expected losses derived from the selected pure premiums for the July 1, 1938 rate revision because it was believed

* The actual losses were not taken from Schedule Z for 1938 but rather from these exhibits because in the latter losses are conveniently grouped by Serious, Non-Serious and Medical. For classifications involving ex-medical coverage, the medical losses shown in the classification experience exhibits are increased to a statutory medical coverage basis, however. In order to eliminate the effect of this extrapolation, actual medical losses modified by the factors applied in preparing the classification experience were substituted for the medical losses shown in these exhibits.

that the experience incurred during this period should be expected to accord best with the theoretical experience assumed in selecting these pure premiums. It may, however, be better for the sensitiveness of our test method to compare expected losses with the actual losses of a period comprising two or three policy years, because for a longer period the dispersion of the frequency distribution of x_a will be smaller and, hence, σ_a^2 will be reduced whereas x_n and σ_n^2 will not be affected. The portion x_n of $x = x_a + x_n$ in which we are primarily interested will therefore have greater relative weight and this will enhance the sensitiveness of our tests. In this study only one policy year of actual experience was used because otherwise only the selected pure premiums for a less recent rate revision could have been tested, and also because of the prohibitive volume of the calculations required which is too great to be handled by a single person.

Since only classification relativities are the subject of the present study all law amendment factors, projections factors and rate level change factors were eliminated by multiplying actual losses by adjustment factors which were determined separately for each industry group and each partial pure premium and designed to produce the same aggregate amount of actual and expected losses. These factors were obtained by adding serious, non-serious and medical expected and unadjusted actual losses separately for each industry group and dividing each total of expected losses by the corresponding total of actual losses. A rough check of these adjustment factors was made in the following manner: The actual policy year 1938 losses have been used to determine the level of the selected pure premiums used in the July 1, 1941 rate revision in New York. The rate level change factor which translates July 1, 1938 rates to the level of the July 1, 1941 selected pure premiums is known and so are the factors which translate the July 1, 1938 selected pure premiums to the July 1, 1938 rate level. For this reason theoretical factors could be computed which translate July 1, 1938 selected pure premiums by industry group to policy year 1938 actual losses and these factors agreed reasonably closely with the factors obtained in the manner mentioned before. Throughout this study, the premiums for Federal classifications were split up into two independent parts, one reflecting New York coverage and the other reflecting coverage under the

United States Longshoremen's and Harbor Workers' Act, and separate adjustment factors were computed for both coverages.

The quantity

$$(1') \qquad x = \log_{10}\left[\frac{\text{actual losses}}{\text{expected losses}} \cdot \text{adjustment factor}\right]$$

was finally adopted as the measure of the deviation of actual from expected losses and was computed to two decimal places, separately for each partial pure premium and each classification. The information thus obtained was recorded on punch cards showing (1) the classification code, (2) the industry group code, (3) the absolute amount of x, (4) a code for the sign of x, (5) a code for the state credibility of the pure premium, and (6) a code showing whether x referred to serious, non-serious or medical pure premiums.

Originally, it had been intended to combine the quantities x for serious, non-serious and medical pure premiums, but a test showed that there exists a strong correlation between the deviations for medical pure premiums and those for either serious or non-serious indemnity pure premiums of the same classification. This result is not surprising, since an abnormally low incidence of accidents involving serious (or non-serious) indemnity losses for any classification will have the corollary effect of reducing the medical losses caused by such accidents and, consequently, of depressing the total amount of medical losses. The basic assumption that every x constituted an independent fortuitous event would, hence, have been incorrect if the deviations for medical pure premiums had been combined with those for indemnity pure premiums. It was therefore decided to examine serious, non-serious and medical pure premiums separately.

Another difficulty arises in connection with the deviations produced by serious indemnity pure premiums. Since, so far as New York is concerned, a serious indemnity loss cannot be small but runs into a considerable amount of money, the ratio of actual to expected serious losses for those classifications which have a very low expected frequency of serious losses, will not be distributed continuously, but rather be either zero or a substantial positive quantity. In these cases, therefore, it is not quite accurate to assume that x has a continuous frequency distribution. A difficulty of a more mathematical nature is due to the fact that actual serious, non-serious or medical losses incurred under certain classifications during 1938 were zero and consequently the logarithm of the ratio of these losses to expected losses was $-\infty$. Means or a standard deviation of frequency distributions including non-vanishing frequencies at the point $x = -\infty$ cannot be computed. For many purposes, however, the values $x = -\infty$ can reasonably be excluded. Thus, in particular, where needed in this study, means and variances were computed from the distribution of those values of x which were not equal to $-\infty$.

It may be of interest, although not strictly connected with the subject of this paper, to mention a device by which the burdensome numerical calculations involved in the computation of variances for numerous different values of x have been simplified with the help of Hollerith tabulating machines of the kind that are generally used by insurance carriers. This is explained in the Appendix of this paper.†

E. *Discussion of the Frequency Distributions of the Deviations of Serious, Non-Serious and Medical Actual Losses from the Corresponding Expected Losses*

On charts I, II and III the frequency distributions of the variate x are shown in intervals of .05, for serious, non-serious and medical pure premiums, separately for pure premiums with state credibilities of 50% and over (broken lines) and for pure premiums with state credibilities under 50% (solid lines). The rectangles shown on the left represent the area which is due to the occurrence of a certain number of values of $x = -\infty$. The dotted line represents a normal distribution fitted to the curve for pure premiums with state credibilities of 50% or over, after the values of $x = -\infty$ have been excluded.

The numerical characteristics of the frequency distribution shown in the three charts are exhibited in the following table. This table is discussed under point (e) below.

† The author acknowledges gratefully the great help he has received from Mr. Daniel Kalish of the Compensation Insurance Rating Board who assisted him by preparing the punch cards and the numerous tabulations required by this study.

TABLE A

Analysis of Charts I, II and III

Distribution	Total No. of Values Used	Distribution of x Excluding Values $x = -\infty$						
		Number of Values Exclusive of Values $x = -\infty$	Mean $\bar{x}$	Standard Deviation σ	Standard Deviation of Mean $(5) \div \sqrt{(3)}$	$\frac{\bar{x}}{(6)}$	Probability of Mean in Excess of $\pm(4)$	Deviation of Mean from Zero Significant
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Chart I—Serious Pure Premiums								
(a) P.P.'s with Credibilities 50% & Over	139	133	−.02579	.24435	.02119	−1.217	.22	no
(b) P.P.'s with Credibilities Under 50%	441	215	+.13074	.39738	.02710	+4.824	<.01	yes
(c) Normal Distribution	133	133	.00000	.24435	—	—	—	—
Chart II—Non-Serious Pure Premiums								
(a) P.P.'s with Credibilities 50% & Over	284	284	−.02972	.20451	.01214	−2.448	.02	yes
(b) P.P.'s with Credibilities Under 50%	295	270	−.09367	.41030	.02497	−3.751	<.01	yes
(c) Normal Distribution	284	284	.00000	.20451	—	—	—	—
Chart III—Medical Pure Premiums								
(a) P.P.'s with Credibilities 50% & Over	298	298	−.01302	.13770	.00798	−1.632	.10	no
(b) P.P.'s with Credibilities Under 50%	282	271	−.05469	.27950	.01698	−3.221	<.01	yes
(c) Normal Distribution	298	298	.00000	.13770	—	—	—	—

As said before, no definite conclusions can be drawn with regard to the accuracy of the present method of selecting pure premiums from an analysis of the information contained in charts I, II and III alone. The following general conclusions can, however, be drawn from such an analysis:

(a) After exclusion of the values $x = -\infty$ which constitute an important item only for serious losses and which occur much more frequently among pure premiums with low credibilities, the curves are fairly symmetrical. This means that a positive deviation in the amount of, say, c is about equally likely as a negative deviation in the amount of $-c$. Since the variate represents the logarithm of the ratio of actual to expected losses this additive symmetry of x corresponds to a multiplicative symmetry of the ratio of actual to expected losses. In other words, it is about equally likely that actual losses will amount to, say, 125% of expected losses as that actual losses will amount to $\frac{1}{125\%} = 80\%$ of actual losses.

This circumstance is somewhat at variance with the general practice in casualty insurance of using arithmetic (weighted or unweighted) averages.

(b) It is evident from a comparison of the actual frequency distribution for classifications with high credibilities with the corresponding normal distribution that the actual distributions are much more peaked than the normal distributions and the deviation from the normal form is so large that it cannot be attributed to mere chance. Indeed, an χ^2 test confirms this fact which can be directly inferred from an inspection of the charts. No attempt was made to adjust the actual distributions by means of a mathematical formula representing a theoretical distribution which is more peaked than the normal distribution because most of the statistical criteria of significance have been developed only for normal or near normal distributions.

(c) The actual distributions for classifications with low credibilities show much larger dispersion than those for classifications with high credibilities on each of the three charts. This fact is not surprising if one considers that the major portion of the variate x is due to the deviation of actual from "true" expected losses and that the classifications with low credibilities having much smaller exposure would naturally produce more widely fluctuating actual losses than classifications with large exposure which themselves may be considered as composed of several units of small exposure.

(d) A comparison of the frequency distributions of the deviations for serious, non-serious and medical pure premiums with high credibilities shows that the distribution for medical pure premiums is much more closely concentrated about the mean than that for non-serious pure premiums, and the distribution for non-serious pure premiums is in turn much more concentrated about the mean than the distribution for serious pure premiums. This suggests that the credibility criteria which were used to segregate the pure premiums with high credibilities are not statistically equivalent measures of exposure for the serious, non-serious and medical pure premiums. It appears that the credibility criteria for medical pure premiums are stricter than those for non-serious pure premiums and these are stricter than the criteria for serious pure premiums. It would be desirable, if the split of pure premiums into serious, non-serious and medical portions is to be retained at all, to devise credibility criteria which are statistically equivalent in the sense that, for pure premiums with equal credibility, actual losses concentrate in the same degree about the expected losses irrespective of whether we deal with serious, non-serious or medical losses.

(e) Since the table shown above is mainly self-explanatory, only an observation with respect to columns (7), (8) and (9) will be made. Although the distributions on charts I, II and III do not follow exactly the pattern of a normal distribution, it can be assumed that the means of these distributions are normally distributed with sufficient approximation as to permit the use of the integral of the normal distribution in estimating the probability that the deviations of the means from zero are as great as shown in column (4) of Table A. By doing this the probabilities in column (8) were obtained and, considering 2% as the level of significance, the conclusions in column (9) were formulated. It appears from this table that the deviation of the mean from zero is significant for all pure premiums with credibilities of less than 50% and also for non-serious pure premiums with credibilities of 50% and over. Since actual losses were modified by a common factor so as to produce an aggregate loss volume equal to that of the expected losses, the significantly negative mean for non-serious and medical pure premiums with credibilities of less than 50% indicates that, on the average, actual losses for pure premiums with low credibilities run somewhat lower than expected losses as compared with pure premiums with high credibilities. Pure premiums with low credibilities include a large portion of national pure premiums; this circumstance would therefore suggest that national pure premiums are somewhat too high, although the national pure premiums, if weighted by the product of payrolls and national credibility, are on the correct level. The significance tests for serious pure premiums with credibility under 50% are not quite conclusive because of the large number of value $x = -\infty$ which were excluded from the computation of the means.

F. *Test of the Relative Accuracy of Formula Pure Premiums Based on National Pure Premiums and of Formula Pure Premiums Based on Underlying Pure Premiums*

In order to give an illustrative application of the method evolved in the foregoing pages, a test has been made to measure the relative accuracy of the selected pure premiums prepared for the July 1, 1938 rate revision which are essentially formula pure premiums computed by weighting the indicated state pure premiums against the corresponding national pure premiums and a corresponding set of formula pure premiums based on weighting the indicated state pure premiums against the underlying pure premiums brought to the same level. This test was made only to illustrate the test method outlined and not because the author believes the present method of determining formula pure premiums should be abandoned in favor of formula pure premiums incorporating underlying pure premiums instead of national pure premiums. Since such a plan‡ has, however, been considered as a possible substitute for the present ratemaking procedure, the author believes that this test may also have some interest beyond that of a mere illustration. The test has been applied to all pure premiums with credibilities of less than 50%, since for pure premiums with higher credibilities the indicated state pure premium is the predominant part and, therefore, the substitution of underlying pure premiums for national pure premiums would have only a slight effect. Pure premiums for classifications which have no state credibility at all for any of the three pure premium parts

‡ See *A. G. Smith, Pure Premiums for Compensation Insurance,* P.C.A.S. Vol. XXIV, pp. 35ff.

were excluded because these classifications consist chiefly of non-reviewed classifications for which the underlying pure premiums are identical with the national pure premiums. The formula pure premiums obtained by weighting the state indication against the underlying pure premiums were used without any modification in order to simplify the rather burdensome numerical computations. If selected pure premiums based on underlying pure premiums should actually be applied, the method would certainly need some modification in order to avoid having the pure premiums for classifications with no, or a very low, state credibility perpetuated.

The results of this test are shown on chart IV. The solid line represents the frequency distribution of x_1 based on national pure premiums and the broken line represents the frequency distribution of x_2 based on underlying pure premiums. The graphs for x_1 differ somewhat from the corresponding graphs for pure premiums with credibilities under 50% shown on charts I, II and III because of the exclusion of all pure premiums of classifications with no state credibility for any part of the pure premium. The numerical characteristics of the distributions described by the graphs shown on chart IV are summarized in the following Table B:

TABLE B

TEST OF SIGNIFICANCE OF THE DIFFERENCE OF MEANS AND VARIANCES FOR THE TWO SETS OF FORMULA PURE PREMIUMS

			Test of Significance of Means							
	Number m of P.P.'s Examined	$\sqrt{m}$	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_1 - \bar{x}_2$	$\sigma^2_{\bar{x}_1-\bar{x}_2} = \sigma_0^2$	$\sigma_{\bar{x}_1-\bar{x}_2} = \dfrac{\sigma_0}{\sqrt{m}}$	$\dfrac{\bar{x}_1 - \bar{x}_2}{\sigma_{\bar{x}_1-\bar{x}_2}}$	Probability that $\bar{x}_1 - \bar{x}_2$ Exceeds $\pm$(6)	Difference of Means Significant (2% Level)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Serious P.P.	204	14.28	+.116667	+.119559	−.002892	.003258	.003997	− .724	.47	no
Non-Serious P.P.	195	13.96	−.047282	−.046205	−.001077	.001513	.002786	− .387	.70	no
Medical P.P.	182	13.49	−.011868	−.006154	−.005714	.001214	.002583	−2.215	.03	no

	Test of Significance of Difference of Variances										
					Assuming $r \leqq .85$						
	σ_1^2	σ_2^2	$\Delta = \lvert\sigma_1^2 - \sigma_2^2\rvert$	$t = \dfrac{\sigma_0^2}{\Delta}$	Lower Limit for s	Upper Limit for s	$\dfrac{\log_e(16)}{1/\sqrt{m}}$	$\dfrac{\log_e(17)}{1/\sqrt{m}}$	Probability that s Exceeds (16) ± 1	Probability that s Exceeds (17) ± 1	Difference Between σ_1 and σ_2 Significant
(1)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
Serious P.P.	.142530	.147300	.004770	.683	1.310	4.053	3.857	19.989	< .01	< .01	yes
Non-Serious P.P.	.092576	.092257	.000319	4.743	1.031	1.534	.417	5.975	.68	< .01	?
Medical P.P.	.041758	.041497	.000261	3.651	1.042	1.325	.555	3.792	.58	< .01	?

	Test of Significance of Difference of Variances						
	No Upper Limit for r Assumed						
	Lower Limit for s	Upper Limit for s	$\dfrac{\log_e(23)}{1/\sqrt{m}}$	$\dfrac{\log_e(24)}{1/\sqrt{m}}$	Probability that s Exceeds (23) ± 1	Probability that s Exceeds (24) ± 1	Difference Between σ_1 and σ_2 Significant
(1)	(23)	(24)	(25)	(26)	(27)	(28)	(29)
Serious P.P.	1.017	5.309	.235	23.844	.81	< .01	?
Non-Serious P.P.	1.002	1.534	.024	5.975	.98	< .01	?
Medical P.P.	1.003	1.325	.042	3.792	.97	< .01	?

Before explaining in detail the meaning of this table it is necessary to digress on the theory of tests of significance as applicable in this case.

The first test consists in determining whether the difference between the means of the distributions of x_1 and those of x_2 is significant. In reality, we are interested in the distribution of $x_{n,1}$ which measures the deviation from the "true" expected losses of the expected losses based on selected pure premiums derived from national pure premiums and in the distribution of $x_{n,2}$ which measures the deviation from "true" expected losses of the expected losses based on selected pure premiums derived from underlying pure premiums. Yet, since the "true" expected losses are unknown, $\bar{x}_{n,1}$ and $\bar{x}_{n,2}$ cannot be computed. It was, however, shown in formula (2) that:

$$(2') \qquad x_1 = x_a + x_{n,1} \qquad\qquad x_2 = x_a + x_{n,2}$$

and consequently

$$(7) \qquad x_{n,1} - x_{n,2} = x_1 - x_2 \quad \text{and} \quad \bar{x}_{n,1} - \bar{x}_{n,2} = \bar{x}_1 - \bar{x}_2$$

where the bar indicates a mean. We can, therefore, test whether $\bar{x}_1 - \bar{x}_2$ is significant. For this purpose, the variance

$$(8) \qquad \sigma_0^2 = \sigma_{x_1-x_2}^2 = \sigma_{x_{n,1}-x_{n,2}}^2$$

was determined and shown in column (7) of table B. If m is the number of values of $(x_1 - x_2)$, which is shown in column (2), the variance of the mean will be $\sigma_{\bar{x}_1-\bar{x}_2}^2 = \sigma_0^2/m$ and the standard deviation $\sigma_0/\sqrt{m}$. This latter is shown in column (8) of Table B. It is assumed that the mean $(\bar{x}_1 - \bar{x}_2)$ is distributed normally with sufficient approximation to permit the use of a table of the integral of the normal distribution in order to determine the probability that the mean of $(x_1 - x_2)$ exceeds the amount shown in column (6). This is done in columns (10) and (11) and it is found that none of the differences between means is significant. In other words neither of the distributions of $x_{n,1}$ and $x_{n,2}$ contains a systematical bias with respect to the other.

Having thus satisfied ourselves that the two distributions have about the same mean we proceed with a study of the variances. In columns (12) and (13) of Table B are shown the variances σ_1^2 and σ_2^2 of x_1 and x_2 and in column (14) the absolute amount of the difference $\Delta = \sigma_1^2 - \sigma_2^2$. The latter is negative for serious pure premiums and positive for non-serious and medical pure premiums. Since

$$(6') \qquad \sigma_1^2 = \sigma_a^2 + \sigma_{n,1}^2 \qquad\qquad \sigma_2^2 = \sigma_a^2 + \sigma_{n,2}^2$$

this would indicate that

$\sigma_{n,1}^2 < \sigma_{n,2}^2$ for serious pure premiums

$\sigma_{n,1}^2 > \sigma_{n,2}^2$ for non-serious and medical pure premiums

Or, in words, the formula pure premiums based on national pure premiums are more accurate for serious pure premiums and the formula pure premiums based on underlying pure premiums are more accurate for non-serious and medical pure premiums. Before making such a statement, however, it is necessary to determine, by means of a test of significance, whether these results are not perhaps merely due to chance.

It is shown in the theory of the "Analysis of Variances"§ that the expression

$$\log_e \frac{\sigma}{\sigma'}$$

for two samples with variances σ and σ' and consisting of m and m' elements respectively and whose means are not significantly different is about normally distributed with mean *0* and standard deviation $\sqrt{\frac{1}{2}\left(\frac{1}{m} + \frac{1}{m'}\right)}$, if the samples are supposed to be derived from a statistical population whose distribution is not too different from the normal distribution. If one wishes to determine whether the two variances do not differ from each other significantly or, in other words, whether the two samples may be considered as derived from the same parent population, one calculates the probability that $\log_e \frac{\sigma}{\sigma'}$ exceeds the observed value (using a table of the integral of the normal frequency distribution). If this probability is larger than the adopted level of significance, usually .05 or .02, then the observed difference in the variances is considered not significant and merely due to chance.

In our case we would have to consider the expression

$$(9) \qquad \log_e \frac{\sigma_{n,1}}{\sigma_{n,2}} = \log_e s, \qquad \text{where } s = \frac{\sigma_{n,1}}{\sigma_{n,2}}$$

§ See *R. A. Fisher, Statistical Methods for Research Workers,* 4th ed., Edinburgh and London 1932, pp. 206 ff.

which would be about normally distributed with a standard deviation of $\sqrt{\frac{1}{m}}$ where $m = m'$ is the number of pure premiums used which is shown in column (2) of Table B. Unfortunately we do not know $\sigma_{n,1}$ and $\sigma_{n,2}$ and must, therefore, estimate the expression $\log_e s$ on the basis of the available data. We shall try to obtain a lower and an upper limit for $\log_e s$. If then a test reveals that even the upper limit for $\log_e s$ is not significant or that already the lower limit is significant, we can draw a definite conclusion regarding the significance of $\log_e s$. In the present case the limits are, unfortunately, so wide apart that the lower limit is not significant and the upper limit is significant so that no final conclusion can be drawn.

We start by developing the general theory. From the actual experience we determine

(10) $\sigma_1^2 = \frac{1}{m} \Sigma\, (x_1 - \bar{x}_1)^2 = \sigma_a^2 + \sigma_{n,1}^2$

(11) $\sigma_2^2 = \frac{1}{m} \Sigma\, (x_2 - \bar{x}_2)^2 = \sigma_a^2 + \sigma_{n,2}^2$

(12) $\triangle = \sigma_1^2 - \sigma_2^2 = \sigma_{n,1}^2 - \sigma_{n,2}^2$

(13) $\sigma_0^2 = \frac{1}{m} \Sigma\, [\,(x_{n,1} - \bar{x}_{n,1}) - (x_{n,2} - \bar{x}_{n,2})\,]^2$

$$= \frac{1}{m} \Sigma\, (x_{n,1} - \bar{x}_{n,1})^2 - \frac{2}{m} \Sigma\, (x_{n,1} - \bar{x}_{n,1})(x_{n,2} - \bar{x}_{n,2}) + \frac{1}{m} \Sigma (x_{n,2} - \bar{x}_{n,2})^2$$

If $r = \dfrac{\Sigma\, (x_{n,1} - \bar{x}_{n,1})(x_{n,2} - \bar{x}_{n,2})}{\sigma_{n,1} \cdot \sigma_{n,2}}$ designates the correlation coefficient of $x_{n,1}$ and $x_{n,2}$ this can be written

(14) $$\sigma_0^2 = \sigma_{n,1}^2 - 2\, r\, \sigma_{n,1}\, \sigma_{n,2} + \sigma_{n,2}^2$$

r is a number which can vary from -1 to $+1$, but usually some plausible assumption regarding r can be made. If we write

(15) $$t = \frac{\sigma_0^2}{\triangle}$$

we can express s by means of the known quantity t and the estimated quantity r in the following manner:

$$(16) \qquad t = \frac{\sigma_{n,1}^2 - 2\,r\,\sigma_{n,1}\,\sigma_{n,2} + \sigma_{n,2}^2}{\sigma_{n,1}^2 - \sigma_{n,2}^2} = \frac{s^2 - 2\,r\,s + 1}{s^2 - 1}$$

and, hence,

$$(17) \qquad s = -\frac{r}{t-1} \pm \frac{1}{t-1}\sqrt{r^2 + t^2 - 1}$$

Let us, for the sake of simplicity, assign the subscript 1 to that distribution for which $\sigma_{n,1}^2 > \sigma_{n,2}^2$ then Δ and, consequently, t will be positive. Since s is essentially a real and positive number, we derive from the above relation:

(a) If $t < 1$

(18) $r^2 > 1 - t^2$ and hence $r > +\sqrt{1 - t^2}$

since, for $r < 0$, t would be > 1.

(b) If $t > 1$ only the $+$ sign before the square root provides a suitable solution as, otherwise s becomes negative.

If the two sets of pure premiums are based on entirely different principles, we may assume that the correlation coefficient r of the deviations of the two sets of pure premiums from the "true" pure premiums and is zero. In this case

$$\sigma_0^2 = \sigma_{n,1}^2 + \sigma_{n,2}^2 > \Delta = \sigma_{n,1}^2 - \sigma_{n,2}^2$$

and consequently

$$(19) \qquad t > 1 \quad s = +\frac{1}{t-1}\sqrt{t^2 - 1} = +\sqrt{\frac{t+1}{t-1}}$$

and we will be able to apply the test of significance to the natural logarithm of this quantity.

It cannot be said, however, that the two particular sets of formula pure premiums under investigation in this study are based on entirely different principles, as both have a certain portion of indicated state pure premiums in common and also because the underlying pure premiums are not entirely independent of the national pure premiums. We can, therefore, only assume that the correlation coefficient r is non-negative and, probably, not too near to unity. Let us, therefore, assume

(20) $\sqrt{1-t^2} \leqq r \leqq .85$ for $t < 1$ and $0 \leqq r \leqq .85$ for $t > 1$

then it can easily be shown that,

(a) if $t < 1$

(21) $$\frac{.85}{1-t} - \frac{1}{1-t}\sqrt{t^2 - .2775} \leq s \leq \frac{.85}{1-t} + \frac{1}{1-t}\sqrt{t^2 - .2775}$$

(b) if $t > 1$

(22) $$-\frac{.85}{t-1} + \frac{1}{t-1}\sqrt{t^2 - .2775} \leq s \leq \sqrt{\frac{t+1}{t-1}}$$

If we do not assume the probable, but somewhat arbitrary, upper limit of $+.85$ for r, the limits for s would be

(a) if $t < 1$

(23) $$\frac{\sigma_1}{\sigma_2} \leq s \leq \frac{1+t}{1-t}$$

(b) if $t > 1$

(24) $$\frac{\sigma_1}{\sigma_2} \leqq s \leqq \sqrt{\frac{t+1}{t-1}}$$

The lower and upper limits for s are shown in columns (16), (17), (23) and (24), separately for an assumed ceiling of .85 for r and for an unlimited r. The probabilities that $s = \frac{\sigma_{n,1}}{\sigma_{n,2}}$ exceeds these limits are shown in columns (20), (21) and (27), (28) respectively. It follows that, generally, the lower limits are not significant and the upper limits are, so that, the actual value of $s = \frac{\sigma_{n,1}}{\sigma_{n,2}}$, being somewhere in between, no definite conclusion can be drawn. Only the difference in the variances for serious pure premiums is significant, under the assumption $r \leqq .85$, even if $\frac{\sigma_{n,1}}{\sigma_{n,2}}$ does not surpass its lower limit. A definite answer can be obtained only if we assume that the correlation coefficient r is near its ceiling—an assumption which seems rather probable. In this case, we should have to conclude that formula pure premiums based on underlying pure premiums furnish a better fit than those based on national pure premiums for non-serious and medical pure premiums and a worse fit for serious pure premiums. The author has been at a loss to explain this difference in the behavior of the various types of pure premiums, unless the exclusion of a large number of values $x = -\infty$ and the previously discussed discontinuity of actual losses for small classifications has the effect of distorting the results for the serious pure premium group.

The difficulties encountered in testing the significance of the relative accuracy of selected pure premiums determined by the present ratemaking procedure and of formula pure premiums based on underlying pure premiums are not likely to recur if present selected pure premiums are compared with a set of pure premiums which are not based on similar principles. The author hopes to develop such a set of pure premiums in a subsequent part of this paper.

APPENDIX

A Method of Computing Sums of Products with the Help of Non-Multiplying Tabulating Machines

The method to be discussed is applicable wherever the total of the products of two different or identical sets of factors is to be computed, although a knowledge of the value of the individual products is not required. This type of problem occurs not only in statistics when variances or co-variances have to be determined, but also in casualty insurance wherever a set of payrolls has to be multiplied by a set of rates in order to determine the aggregate premium figure.

Let the first set of factors be A_1, A_2, $\cdots$, A_i, $\cdots$, and the corresponding second set of factors of the form $100\,a_1+10\,b_1+c_1$, $\cdots$, $100\,a_i+10\,b_i+c_i$, $\cdots$ where a_i, b_i, c_i, are each one of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. The sum $\Sigma\,[A_i\,x\,(100\,a_i+10\,b_i+c_i)]$ can be written $100\,\Sigma\,A_i\,a_i+10\,\Sigma\,A_i\,b_i+\Sigma\,A_i\,c_i$ and each of the three subsums can itself be split into ten minor totals, each of which has the same digit as a second factor and can, therefore, be written as a product of this digit with the corresponding total of the A_i. The use of tabulators is based on this resolution of the original total.

Each pair of factors A_i and $(100a_i + 10b_i + c_i)$ is punched on one punch card. These cards are sorted on the first digit a_i of the second factor and subtotals of A_i determined for each of the ten different digits a_i. In the same manner the subtotals corresponding to the different digits b_i and c_i are determined. The only computation necessary is the multiplication of the subtotals by the corresponding digits and that power of 10 which indicates the place of this digit and the totaling of all products thus obtained. The calculations are thus reduced to at the most twenty-seven multiplications with a one-digit factor (if the second factor has not more than three digits). This procedure saves a great deal of time where the number of products is very numerous.

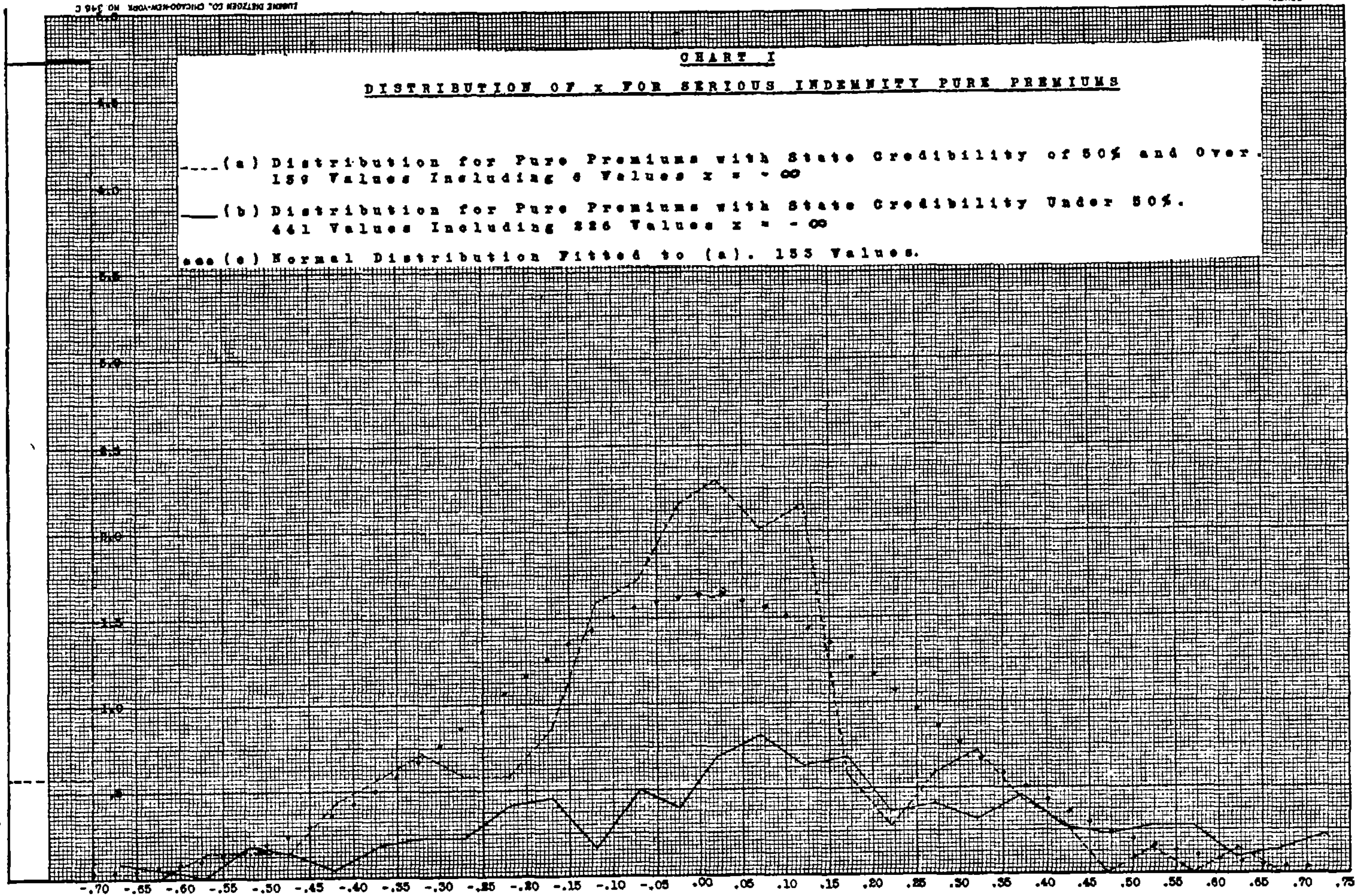
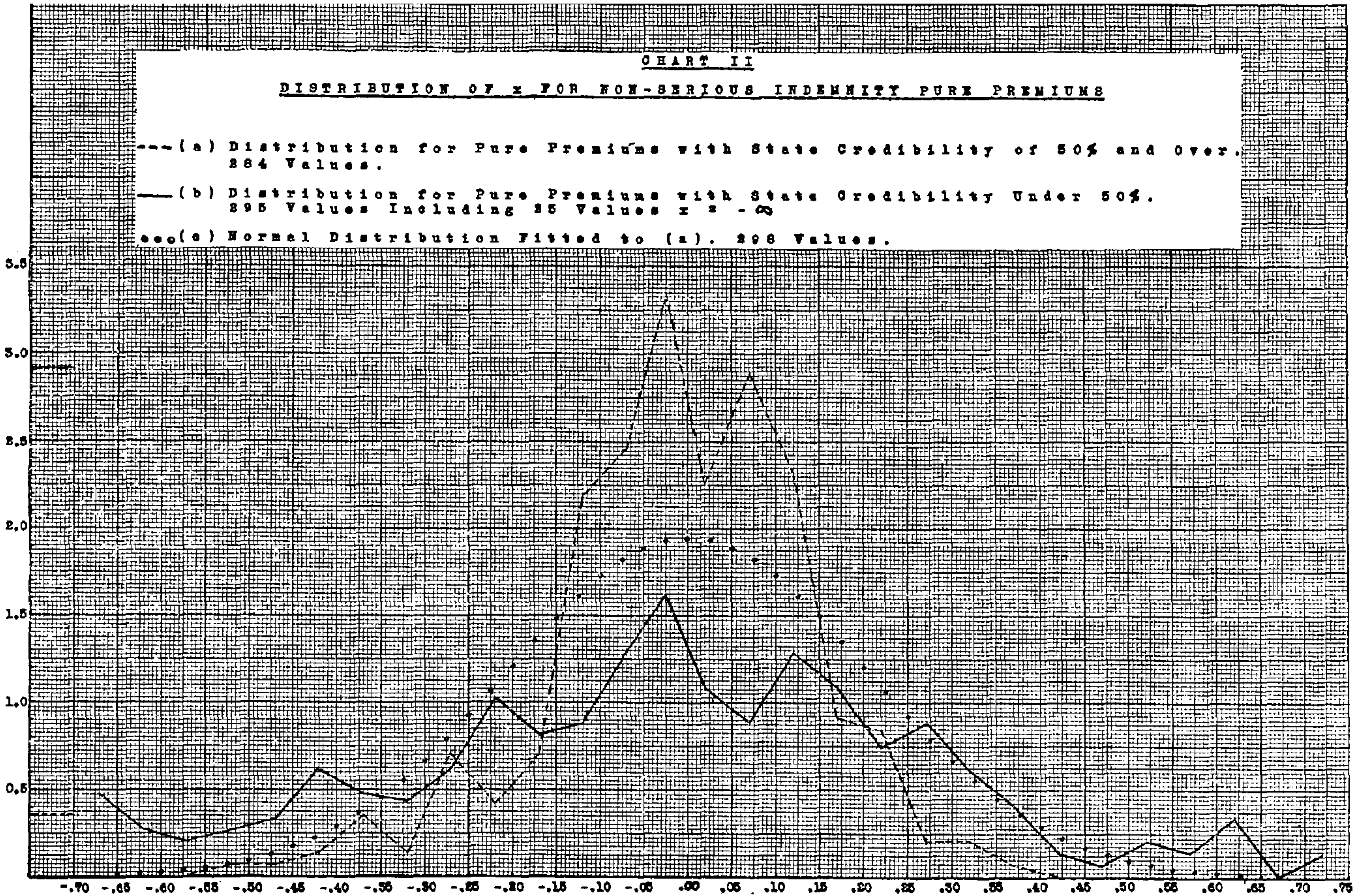


CHART I
DISTRIBUTION OF x FOR SERIOUS INDEMNITY PURE PREMIUMS
(a) Distribution for Pure Premiums with State Credibility of 50% and Over.
159 Values Including 6 Values x = - ∞
(b) Distribution for Pure Premiums with State Credibility Under 50%.
441 Values Including 226 Values x = - ∞
(c) Normal Distribution Fitted to (a). 153 Values.
1.0
1.5
2.0
-.70 -.65 -.60 -.55 -.50 -.45 -.40 -.35 -.30 -.25 -.20 -.15 -.10 -.05 .00 .05 .10 .15 .20 .25 .30 .35 .40 .45 .50 .55 .60 .65 .70 .75

CHART II

DISTRIBUTION OF x FOR NON-SERIOUS INDEMNITY PURE PREMIUMS



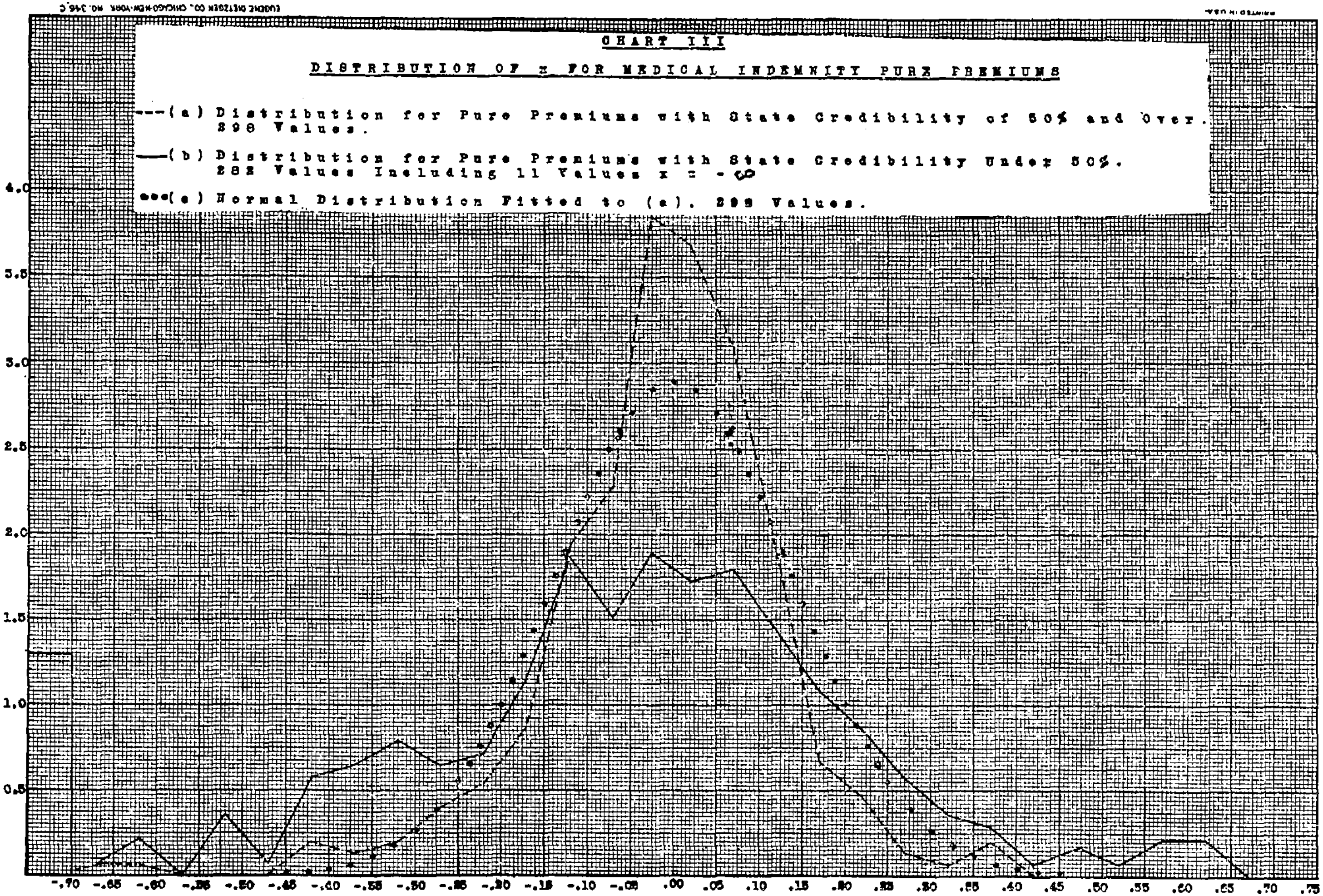


CHART III
DISTRIBUTION OF x FOR MEDICAL INDEMNITY PURE PREMIUMS
---(a) Distribution for Pure Premiums with State Credibility of 50% and Over. 298 Values.
—(b) Distribution for Pure Premiums with State Credibility Under 50%. 282 Values Including 11 Values x = -∞
•••(c) Normal Distribution Fitted to (a). 298 Values.
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
-.70 -.65 -.60 -.55 -.50 -.45 -.40 -.35 -.30 -.25 -.20 -.15 -.10 -.05 .00 .05 .10 .15 .20 .25 .30 .35 .40 .45 .50 .55 .60 .65 .70 .75

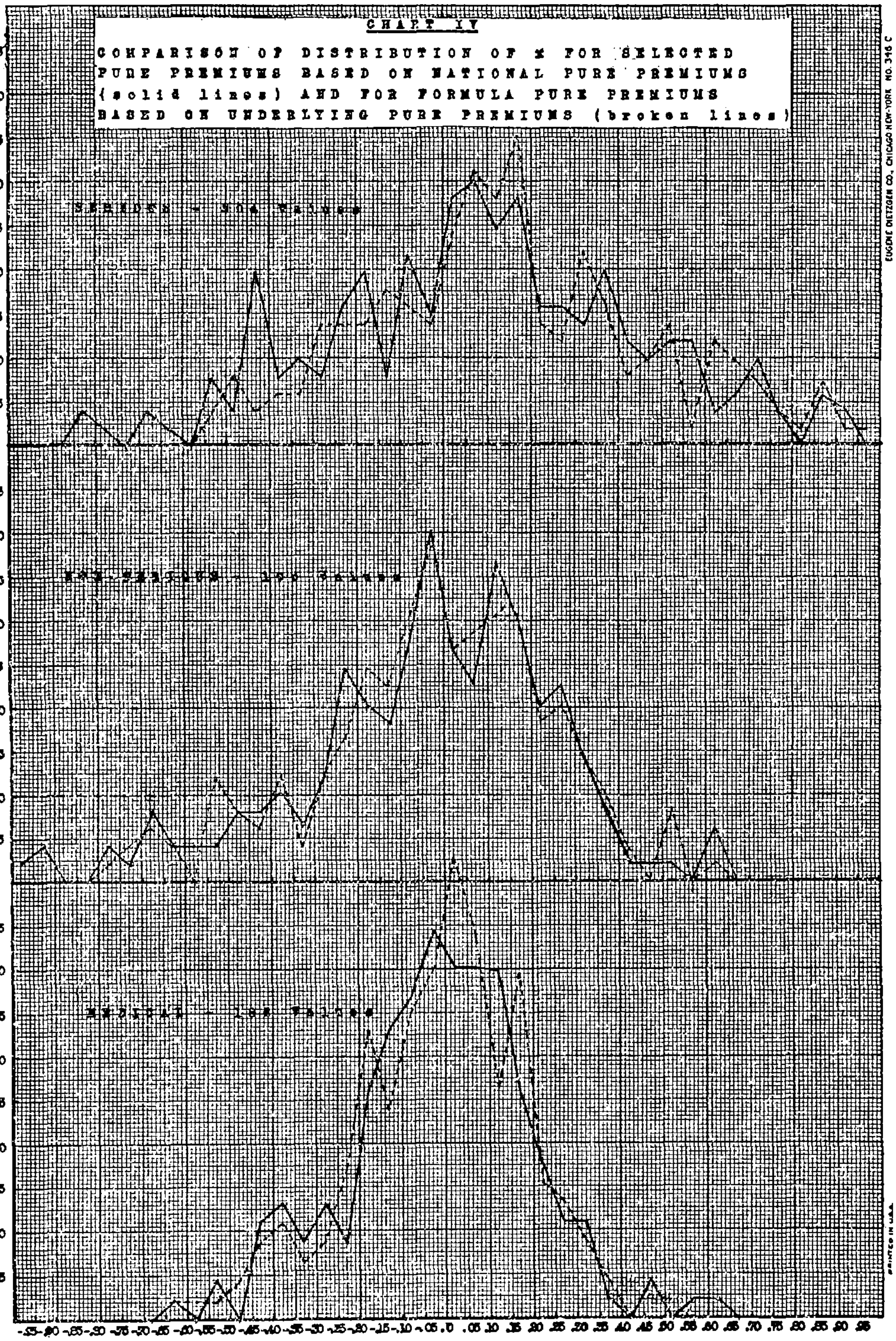


CHART IV
COMPARISON OF DISTRIBUTION OF x FOR SELECTED PURE PREMIUMS BASED ON NATIONAL PURE PREMIUMS (solid lines) AND FOR FORMULA PURE PREMIUMS BASED ON UNDERLYING PURE PREMIUMS (broken lines)
2.25
2.00
1.75
1.50
1.25
1.00
.75
.50
.25
-.95 -.90 -.85 -.80 -.75 -.70 -.65 -.60 -.55 -.50 -.45 -.40 -.35 -.30 -.25 -.20 -.15 -.10 -.05 .0 .05 .10 .15 .20 .25 .30 .35 .40 .45 .50 .55 .60 .65 .70 .75 .80 .85 .90 .95
EUGENE DIETZGEN CO. CHICAGO NEW-YORK NO. 346 C
PRINTED IN U.S.A.