# ABSTRACT OF THE DISCUSSION OF THE PAPERS READ AT THE PREVIOUS MEETING.

## A NEW CRITERION OF ADEQUACY OF EXPOSURE—ALBERT H. MOWBRAY.

VOL. IV, PAGE 263.

WRITTEN DISCUSSION.

MR. WILLIAM LESLIE:

This paper deals with the application of the Tchebycheffian criterion to empirical probabilities of hazard derived from workmen's compensation data. It presents a practical method of analyzing by means of this criterion the data used in constructing compensation rates.

The value of mathematical tests to supplant, or at least supplement, the judgment now used in combining experience and deriving pure premiums for compensation insurance is evident to anyone familiar with either the present method of rate-making or the many inconsistencies found in applying the rates appearing in our compensation manuals.

The particular test here described is one for measuring the payroll exposure necessary to give a predetermined degree of accuracy in the pure premium or, with a given exposure, for calculating the degree of accuracy in the resulting pure premium. As Mr. Mowbray points out, however, inadequacy of exposure is only one of the elements entering into the present necessity for the use of judgment in rate-making. Even were this not so the test here described would not do away with the necessity of using judgment in grouping classifications and erecting pure premiums where the exposures are inadequate. Its value lies in the guide it affords in the use of judgment, both in furnishing a measure of the extent of departure from the indicated pure premium warranted by the paucity of the data and also in giving the exposure which should be obtained by appropriate grouping of classifications to give dependable pure premiums. If it is based upon sound principles, it should prove of great assistance in rate-making.

The weak point in the application of this criterion to workmen's compensation experience lies in the assumption that the hazard probabilities constitute a Bernoullian Series. Mr. Mowbray recognizes this weakness, as witness the following quotation: " . . . after due allowance and correction has been made for disturbing factors (such as 'increasing cost,' 'industrial activity,' and other items

considered at the last rate conference) the first may perhaps be taken as approximately true. Indeed, its approximate truth when so corrected seems to be a fundamental requirement of prospective rate-making." In this statement "the first" refers to the assumption regarding the constancy of the probability of hazard throughout the period observed. The formula derived by Mr. Mowbray is a development of the application of the criterion of Tchebycheff to a Bernoullian series and we cannot, in the absence of proof to the contrary, assume that it would give even approximately accurate results if the probability of hazard should not remain constant.

In the absence of any knowledge of the true facts, and in the light of the accepted methods of making rates for compensation insurance on the basis of the ratio of total combined losses (with appropriate adjustments) to total combined payroll, we might be justified in accepting the approximate truth of the assumption of a Bernoullian series. It seems to me, however, that further investigation on this point is warranted before applying the criterion in practice.

Arne Fisher in his "Outline of a Method for Determining Basic Pure Premiums," reported in the *Proceedings,* Vol. II, p. 394, casts a doubt upon the identity between a Bernoullian distribution of occurrences and a workmen's compensation loss series. He suggests an actual test of the stability of the series of losses for each classification, but under present conditions of rate-making, such a test of the stability of the pure premiums is impractical.

Albert W. Whitney in his article "The Theory of Experience Rating," published in the *Proceedings,* Vol. IV, p. 274, deals with the theory that each risk within a particular classification has its own real risk hazard and that these risk hazards group themselves about the real class hazard according to some law of frequency. If this is a true hypothesis then the series is not a Bernoulian series because the probability of hazard is not constant but fluctuates from risk to risk within the classification.

Not only does our practical experience teach us that there is very probably a fluctuation in the probability hazard from risk to risk within a given classification, but my personal observation of certain large risks, over a period of several years, indicates to me that the risk hazard fluctuates from year to year, due to changing conditions for which as yet appropriate modifications have not been established.

Perhaps Mr. Mowbray contemplates a reduction to the Bernoullian series by means of judgment corrections applied to cases showing pronounced and evident variations in the risk hazards, but it seems to me that such corrections can at most affect but a relatively small number of risks and if the probability of hazard to begin with is not approximately constant for all risks within the class, then these more or less sporadic adjustments will not convert the data into a Bernoullian series.

These facts indicate to me that the experience for the majority of our manual classifications probably does not embrace a constant probability throughout the period observed. Furthermore I cannot entirely agree with Mr. Mowbray that such a conclusion is inconsistent with the theory of our present system of prospective rate-making. In this connection I believe we must assume that the modifications for increasing cost, etc., as well as the various state differentials have been accurately determined so that our "sets" or "samples" of experience (each constituting the experience under policies of a particular year of issue in a particular state covering various risks falling within the same classification), when corrected and reduced, each contain the same average probability of hazard. If there is no variation within each set then the experience would constitute a Poisson series. In our scheme of prospective rate-making we form a mean or average classification rate from past experience which we hope, in view of the above referred to modification factors, will properly represent the average rate for the classification under consideration. From the theory of probabilities, we know that the mean of a Poisson series with varying probabilities is equal to the mean of a Bernoullian series whose constant probability is the arithmetic mean of the varying probabilities of the corresponding Poisson series. Therefore regardless of whether or not our loss data represents a Bernoullian or a Poisson series, the present method gives a true prospective class rate upon the assumption that the modification factors and the law differentials are accurate.

In applying this class rate to individual risks the several schemes which have been adopted to adjust it to the varying hazards of the risks within the class seem to give further evidence of the varying risk hazard and to me seem to be a tacit admission that the loss series is not normal.

A very interesting feature of Mr. Mowbray's paper is his adoption of the pure premium as a basis for a probability. That it simplifies the procedure in applying the criterion is quite evident. Whether or not it is logical is not quite so clear.

The probabilities dealt with in the paper, and generally met with in practice in connection with compensation hazards, are quite small. Suppose, however, the following purely hypothetical case existed:

Out of each thousand soldiers engaged in the war, seventy die from one cause or another during the year and an average death benefit of nine thousand dollars is paid. If the average annual earnings of each soldier are six hundred dollars, what is the probability that 100 per cent. of the payroll unit will be required for death claims arising out of the expenditure of that unit. It is readily seen that the value of the probability exceeds unity, although the probability of death is only seven one-hundredths.

If Mr. Mowbray's expression for a probability is a logical one,

how is a result such as the above explained? If each of the events in the denominator is equally likely and they constitute the whole range of possible events how can the above situation arise?

MR. ARNE FISHER:

Mr. Mowbray is one of the few actuaries in this country who has made an attempt to extend the application of mathematical analysis to actuarial and statistical work beyond the common rules and comparatively elementary methods usually employed by the actuaries of the life assurance companies. I think that it will be admitted that invalidity and sickness assurance require more refined mathematical methods than those required in ordinary life assurance calculations, and the statisticians and actuaries of our society need a far more extensive mathematical training than that usually attained by an actuary of a life company.

It is, therefore, very pleasing to note that Mr. Mowbray in this article strays away from the well-beaten paths of his colleagues in the life branch, several of whom are, as I once upon a time asserted, suffering from what the Italian philosopher, Morselli, has called "the sterilization of the mind."

Taken as a whole, I agree with the conclusions of the author, but a few points may perhaps be subjected to a more critical analysis. Mr. Mowbray makes frequent use of the term "homogeneous," without defining what he means by homogeneity in statistical series and mass phenomena. As far as I can judge, his classification of statistical data is rather of a subjective kind. What Mr. Mowbray would call homogeneous another statistician might indeed consider as heterogeneous. Homogeneity, if such a thing can be said to exist in statistical observations, is not a fixed and universal notion, but is a varying element in itself, since it is evident that there are various degrees of homogeneity. We might, for instance, ask whether the Mongolian race is more homogeneous than the White race. All statistical analysis is in its last instance simply a study of variation, this latter word taken in its most general sense. The majority of statistical mass phenomena exhibit a tendency to cluster around certain norms. But this clustering tendency varies with the statistical object. An interesting example is offered in anthropometric measurements. Measurements of recruits from various countries show decided variations of clustering tendencies around certain norms, as for instance around the mean value. And even inside each locality we find great variations. Probably one would term the measurements of recruits from a snug little country like Denmark as homogeneous. But, strictly speaking, this is not the case. If we take the members of the regiment of the Royal Guards, none of whom are below 6 ft. 2 in. in height, we will find a much denser clustering around the mean than in the case of the other arms of the service. Unless we are able to express this varying degree of homogeneity by means of abstract numbers, the volu-

minous verbalisms about "homogeneous" material, of which actuaries and statisticians are so fond of speaking, becomes to a very large extent a matter of personal judgment.

Personally, I prefer the term "stability" to that of "homogeneity." This is also a variate, and the question arises as to how we shall measure this varying degree of stability. One of the simplest and speediest test is by the so-called Lexian-Charlier dispersion theory. To quote from Charlier, "when a statistical *homograde* series is given (as, for instance, the rates in compensation insurance) the first task of the statistician is to compare it with the series which in the given case should follow as the consequence of the Bernoullian Theorem. If the series agrees with this theorem it demands, beyond the determination of errors due to random sampling, no *explanation is necessary*—as little as it is necessary to 'explain' why in coin-tossing a head and not a tail appears in a certain case."

I should have wished that Mr. Mowbray had made use of this simple and quick method to test whether the series with which he is dealing are Bernoullian series or not. As it is, we are—as far as I can see—forced to rely upon Mr. Mowbray's personal subjective judgment, which, no matter how excellent and keen it may be, nevertheless does not come up to the exactness of the cold and impersonal analysis by purely objective methods.

Assuming, however, for the moment that Mr. Mowbray's personal subjective judgment is so keen that it can be substituted for the more careful and conservative methods of the objective analysis, we shall, in the absence of further information about the specific details of the data, also assume that the series with which he is dealing are Bernoullian series. Mr. Mowbray now makes use of the criterion of Tchebycheff to test the probability of deviations from the indicated empirical rates. Mr. Mowbray could have simplified his calculation somewhat if he had used the formulas on pp. 110–111 of my book on "Probabilities" instead of the formulas on page 108. As a matter of fact, I have on p. 111 used Tchebycheff's criterion to prove the Bernoullian Theorem. However, this is a mere matter of taste and does not alter the final results as reached by the author. Moreover, the application of Tchebycheff's criterion is very conservative, inasmuch as it over-estimates by a wide margin the limits inside which the expected deviation may occur. The criterion, at least in the form used by Mowbray, however, does not give us the means to determine the probability of the occurrence of a specific deviation.

I trust, therefore, that Mr. Mowbray will pardon me for showing how such probabilities can be shown in tabular form as frequency functions and also in graphical form by means of frequency curves, which, of course, only are plotted from the computed values of the various probabilities. Space forbids me to give the theory and the necessary formula from the theory of the frequency curves of homo-

grade statistics, and I shall limit myself to giving a few actual numerical results relating to the statistical data on p. 270, as furnished by Mr. Cogswell.

The contingencies in question are, with the possible exception of certain temporary disability benefits, events which in the language of mathematical statistics are termed as "rare events," or events whose probability of occurrence is small, *i. e.*, less than, say, .005. Hence we are, unless we had an infinitely large payroll exposure, dealing with what is known as the "Law of Small Numbers." This law, so termed by the Russian statistician Bortkiewicz, was originally introduced by the French mathematician, Poisson, and has of late years been extended and perfected to a very high degree by the members of the modern Scandinavian school of mathematical statistics. Its importance is especially in the theory of risk, much greater than that anachronism which a lot of actuaries and statisticians usually call "the law of averages," a vague and nebulous product of the brains of some academicians, and which, like the ghost in "Hamlet," stalks through the majority of actuarial and statistical writings in this country. The "Law of Small Numbers" is represented by frequency curves of the Poisson-Charlier or the Poisson-Jörgensen type.

As the first illustration I take the death losses in classification 2286. Following the procedure by Mr. Mowbray, let us take $3,000 as the unit of a death loss. We then have:

$s$ = sample set exposed = 13,198 units equivalent to about $39,600,000.

$m$ = number of observed attributes in the sample = 2.364 units equivalent to $7,091.

$q$ = indicated probability or statistical frequency = .000179.

We might now ask: What is the probability to obtain, say, $x$ favorable events in a second sample of the same size? Or, stated in a slightly different form: What is the frequency curve, $F(x)$, of this sample?

This Poisson-Charlier cure is of the form:

$$F(x) = \psi(x) + B_2\Delta^2\psi(x) + B_3\Delta^3\psi(x) + \cdots$$

which has the important property to vanish for all negative values of $x$.

I give below the numerical values of this curve.

$X$ means here the amount of a loss expressed in units of $3,000, and $F(x)$, which is a function of $x$, is the probability of the occurrence of such a loss.

By means of the well-known Gamma Functions it is also possible to interpolate values in this table.* Suppose we wanted to find

* The ordinary interpolation formulas as based upon the finite difference formula of Newton fail to render service here.

the probability that a loss would fall between, say, \$5,400 and \$6,000. An actual computation or interpolation by means of Gamma Functions gives us .0526908 as the probability of the occurrence of such a loss. Likewise we would find a value equal to .0408415 as the probability of the occurrence of a loss between \$9,000 and \$9,600. Similar interpolations can be carried out for arbitrary values of the different losses.

| Loses. | $x$. | $F(x)$. |
|---|---|---|
| Below \$1,500 | 0 | .094043 |
| \$1,500– 4,500 | 1 | .222318 |
| 4,500– 7,500 | 2 | .262780 |
| 7,500–10,500 | 3 | .207072 |
| 10,500–13,500 | 4 | .122379 |
| 13,500–16,500 | 5 | .057861 |
| 16,500–19,500 | 6 | .022797 |
| 19,500–22,500 | 7 | .007699 |
| 22,500–25,500 | 8 | .002275 |
| 25,500–28,500 | 9 | .000598 |
| 28,500–31,500 | 10 | .000141 |
| 31,500–34,500 | 11 | .000030 |
| 34,500–37,500 | 12 | .000006 |
| 37,500 or more | 13 | .000001 |
| | Sum: | 1.000000 |

The interesting point of the above table is, however, the evidence of marked variations due to random sampling in spite of the comparatively large payroll. The curve is decidedly skew, as is seen from a mere glance of its graph (Fig. 1). If we were to fit the curve to a normal curve with standard deviation or Bernoullian dispersion equal to $\sqrt{spq}$ we would obtain a symmetrical curve. I shall not dwell at a closer discussion of a comparison between these two curves at the present, as I intend to discuss the gross fallacy to fit skew frequency distributions by means of the normal curve in one of the illustrations immediately following.

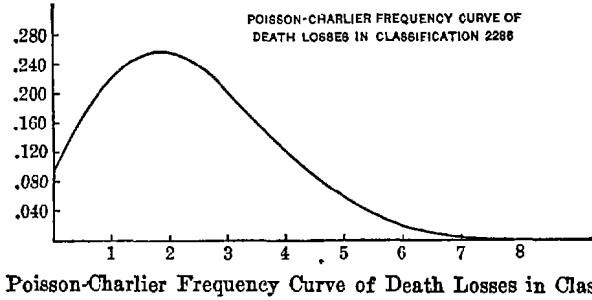


FIG. 1. Poisson-Charlier Frequency Curve of Death Losses in Classification 2286.

Let us in the meantime examine the frequency curve for death losses in classification 2222. We have here, using the same notation as above:

$$m = 8.804$$
$$s = 21{,}781$$
$$q = .000404$$

Our object is to find the frequency curve of this sample. I shall only give the table of $F(x)$ in summary form, as the run of the variations can fully be seen from the graph in Fig. 2. The numerical values are (in groups of 7):

| $x$. | $F(x)$. |
|---|---|
| 0–6 | .225610 |
| 7–13 | .710171 |
| 14–20 | .063890 |
| 20 and over | .000329 |

From the graph it might at a first glance appear that the curve is almost normal in character. This, however, is an optical illusion, due to the fact that the drawing is made on a very small scale. If we consult the actual table we find, however, a decided skewness. This is also seen from the figure where the range to the left of the maximum value or the mode amounts to about 8 intervals or units, while the range to the right of the mode is more than 15 intervals. If the curve had been of the normal type the left half of the range should have been equal to the right half.

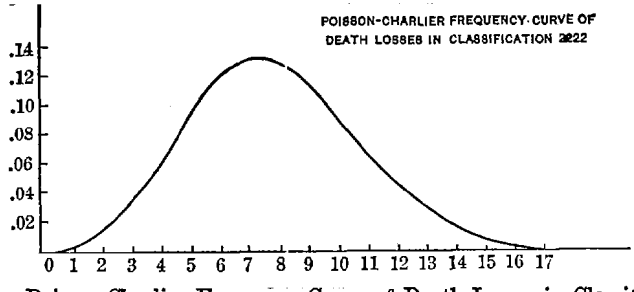


FIG. 2. Poisson-Charlier Frequency Curve of Death Losses in Classification 2222.

As a different illustration I take the permanent total disability losses in classification 2660. I have been informed that the average loss of that kind of invalidity is about $4,000. In order to simplify the computations so as to work with round numbers, I have chosen $3,950 as the unit loss. Whether this be exact or not has no bearing, however, on the construction of the frequency curve. Using this unit we have:

$s = 19{,}985$ (equivalent to a payroll of about \$79,000,000)
$m = 0.9$
$q = .000045$
I give below a table of the frequency function:

| $x$. | $F(x)$. | $x$. | $F(x)$ |
|---|---|---|---|
| 0 | .406570 | 5 | .002001 |
| 1 | .365913 | 6 | .000300 |
| 2 | .164661 | 7 | .000039 |
| 3 | .049398 | 8 | .000002 |
| 4 | .011115 | | |

The graph of the curve is shown in Fig. 3. It is decidedly skew. In fact, it is a one-sided curve.

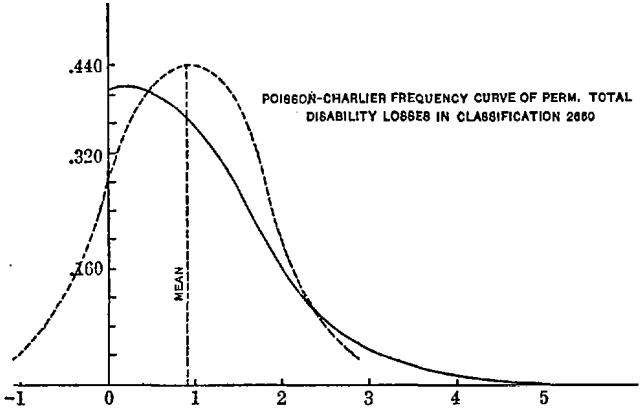


FIG. 3. Poisson-Charlier Frequency Curve of Permanent Total Disability Losses in Classification 2660. Illustrating the error to fit the "Law of Small Numbers" by a Gaussian Normal Error Curve. (Dotted curve represents the Gaussian and full drawn line the Poisson-Charlier Curve.)

We shall now see how the Gaussian normal would fit this frequency distribution. I have fitted the Gaussian curve to the data in order to show the gross fallacy a number of statisticians and actuaries make themselves guilty of in insisting to use this curve. This fallacy has of late been accentuated by the recent articles of Professor Whitney and Mr. Michelbacher on experience rating. If I have understood rightly the trend of these articles, their authors throughout employ the formulas for the Gaussian normal distribution. Let us see how this will turn out in the present case. The standard deviation is in this case given by the formula:

$$\sigma = \sqrt{sqp} = \sqrt{19985 \times .000045 \times 999955} = .94820.$$

The mean or the origin of the curve is equal to $sq = .899325$.

Since only about two thirds of the Gaussian curve is included between the two ordinates corresponding to the abcissas $-\sigma$ and $+\sigma$, that is between the abcissas $-.94820$ and $+.94820$, as reckoned from mean as origin, it is evident that a goodly part of the curve will fall in the region of negative values of the abcissa. This is shown by an actual calculation of the normal curve values as given in the following table:

| Interval. | $F(x)$. |
|---|---|
| Below $-3$ | .00003 |
| From $-3$ to $-2$ | .00110 |
| From $-2$ to $-1$ | .02146 |
| From $-1$ to 0 | .14872 |
| From 0 to 1 | .37066 |
| From 1 to 2 | .33480 |
| From 2 to 3 | .11178 |
| From 3 to 4 | .01090 |
| From 4 to 5 | .00053 |
| From 5 and over | .00002 |

This table shows that 17.121 per cent. of the curve corresponds to negative values of the variate, a fact which is also strikingly illustrated in the graph of the normal curve where the tail piece to the left falls over negative values of the abcissa. This means, of course, the presence of negative losses or actual gains. In other words, the insurer would, if the hypothesis of a normal distribution was true, encounter actual gains from certain death claims. This is too good to be true, and common sense shows the absurdity of such results. In spite of the comparatively large payroll in the class—about $79,000,000—it is evident that on the hypothesis of a normal distribution we encounter a decided negative piece of the tail of the Gaussian curve. It would be of great interest to know what Professor Whitney and Mr. Michelbacher actually have done with this annoying tail-end of the Gaussian curve. Such absurdities do not occur when we use the Poisson-Charlier frequency curve, which, as already stated, has the property to vanish for negative values of the variates.

As a final illustration, I shall take a somewhat different example. Suppose that we were to investigate the variations in the permanent partial disability losses in the classification 2660 on a payroll of $1,000,000. Taking in conformity with Mr. Mowbray $1000 as the unit loss, we have:

$$s = 1000$$
$$q = .000097 \text{ or } .001 \text{ approximately.}$$

This is a decidedly skew distribution, as is seen by a glance from the values of $s$ and $q$. I shall construct the curve for intervals of

losses of $200, that is, 0.2 of the original unit as chosen by Mr. Mowbray. Using the method of interpolation by Gamma Functions we shall have the following values of $F(x)$:

| $x$. | $F(x)$. | $x$. | $F(x)$. |
|---|---|---|---|
| 0.0 | .355389 | 2.2 | .000925 |
| 0.2 | .244218 | 2.4 | .000473 |
| 0.4 | .159460 | 2.6 | .000279 |
| 0.6 | .099909 | 2.8 | .000120 |
| 0.8 | .060470 | 3.0 | .000059 |
| 1.0 | .035530 | 3.2 | .000029 |
| 1.2 | .020352 | 3.4 | .000014 |
| 1.4 | .011390 | 3.6 | .000007 |
| 1.6 | .006244 | 3.8 | .000003 |
| 1.8 | .003360 | 4.0 | .000001 |
| 2.0 | .001777 | | |

Let us again try to see how the Gaussian distribution would fit these results. The mean is here equal to $sq. = 0.1$, while the standard deviation or Bernoullian dispersion amounts to 0.316234. The table of $F(x)$ is as follows:

| $x$. | $F(x)$. | $x$. | $F(x)$. |
|---|---|---|---|
| —1.0 | .0006 | 0.2 | .2399 |
| —0.8 | .0044 | 0.4 | .1610 |
| —0.6 | .0218 | 0.6 | .0723 |
| —0.4 | .0723 | 0.8 | .0218 |
| —0.2 | .1610 | 1.0 | .0044 |
| 0.0 | .2399 | 1.2 | .0006 |

The mode of the normal curve falls at $x = 0.1$. It is not necessary to go into details to show the utter impossibility to attempt to fit this distribution by means of a normal curve, which is also shown at a glance from the graph in Fig. 4.

Almost from the very organization of this Society, I have pleaded for the introduction of more refined mathematical statistical methods in compensation insurance than those commonly used by the life actuaries. I have always insisted that most of the frequency distributions of losses around the average values were essentially skew distributions, because we were dealing with the "Law of Small Numbers." If I had chosen to select my examples from medical statistics relating to inoculation and sample tests on the effect of various vaccines, or if I had selected my numerical examples from the realm of biology, I could easily have shown that my assertions were true. However, it is quite likely that I would have been told that while such things were true in biology they were not necessarily true in compensation insurance. Unfortunately, I did not

have access to any of the statistical data used by the various rate-making experts, so I had to abide my time. Mr. Mowbray and Mr. Cogswell have, however, now supplied me with some authentic data, which has put some very excellent ammunition in my hands, not

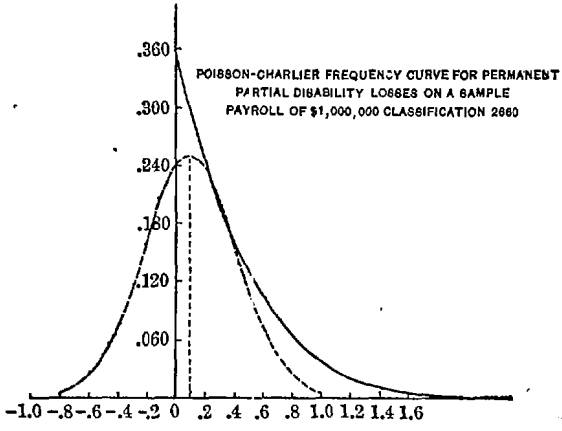


FIG. 4. Poisson-Charlier Frequency Curve for Permanent Partial Disability Losses on a Sample Payroll of $1,000,000 in Classification 2660. Dotted curve represents the fit according to a Gaussian Normal Curve. Full line curve represents the Poisson-Charlier distribution.

only to defend my previous statements, but also to attack some of the methods of my adversaries, if occasion should make this necessary. For this reason I feel very grateful to Mr. Mowbray for having presented his valuable article to the members of this Society.

MR. ALBERT H. MOWBRAY:

(AUTHOR'S REVIEW OF DISCUSSION.)

I am very much gratified to find that my paper has produced two such thoughtful discussions as those presented by Mr. Leslie and Mr. Fisher. I am also pleased to note that both my critics recognize the tentative way in which the results were put forward rather as a pioneer effort in guiding the attempts at solution of the problem of compensation rate-making than a complete solution of the intricate problem discussed.

Mr. Leslie points out that the criterion is applicable only when the hazard probabilities constitute a Bernoullian series and he feels that the hazard probabilities rather constitute a Poisson series. I cannot agree with him in this regard, as it seems to me that probabilities in actual fact very nearly correspond to a Lexian series; but I expressed the view that our various factors employed tended

to reduce the series to a Bernoullian type, on the assumption that such a type was fundamental to our present system of rate-making. To this Mr. Leslie also takes exception, holding that we may apply our present methods equally upon the basis of the probabilities forming a Poisson series, because the mean of a corresponding Poisson and Bernoullian series is the same. I must confess to a much less familiarity with the general dispersion theory than I would like to have, but, if I correctly understand the matter, the fundamental characteristic of a Poisson series as distinguished from a Bernoullian is that in the former the hazard probability is varying and in the latter it is constant. It does not seem to me that the mere fact that in the past the mean of a number of varying probabilities approaches a constant value justifies us in using that for the future unless we assume that that hazard is to remain constant in the future. This, it seems to me, is equivalent to the assumption of a Bernoullian series, and I think that is what Mr. Fisher refers to throughout as determining the question of stability.

Mr. Fisher criticises, and I think with considerable justice from his point of view, my use of the word "homogeneity" in this paper, suggesting in lieu thereof the term "stability." From the mathematical standpoint Mr. Fisher's suggestion is not in the least objectionable, and yet I am not sure but that, from the practical point of view, the term "homogeneity" is preferable. The classification of risks is a problem for underwriters and we have as the first basis a number of manual classifications. It is highly improbable that as applied to the individual risks classified, as, for example, machine shop, Mr. Fisher's tests would show any real homogeneity or stability in the class; yet when the problem of rate-making is under way we have no means of analyzing generally the experience making up a classification. We, of course, are aware that the wire drawing experience in Massachusetts, for example, is dominated if not actually controlled by the experience of the American Steel and Wire Company and that the experience in the electrical apparatus classification in that same state, to use another example, is dominated by the experience of the General Electric Company; but these are rare instances and generally we cannot go back of the classification experience reported to us, and must therefore assume that such experience is homogeneous, perhaps using the term in the way in which we would use it of the Chinese race as compared with other races, notwithstanding that there may be a great variability within the race itself. When, however, we pass beyond the boundaries of the individual classification and combine the experience for several classifications we are taking a further step and the little homogeneity which may be present in the single classification may be further disturbed by the addition of other material, even on the basis of assumed analogy of hazard.

As rate-making is carried on today, both time and money are lacking for mathematical test of stability along the line suggested

by Mr. Fisher. This may be a future development, but, personally, I doubt if we can hope for as much, at least until mathematical methods are better known and mathematical statistics has the appreciation of the public. The executive officers of the companies are not trained in such matters. Heretofore they have entrusted the rate-making to underwriters and rate-making has been wholly a matter of what Mr. Fisher terms subjective judgment. Gradually the actuary has been entrusted with a larger part in the rate-making work, but attempt at too much refinement may tend to deprive the actuary of the position he has so far attained in this respect. It seems to me the need of the hour in many ways is the popularization of mathematical statistics through the preparation of treatises in popular form which will tend to familiarize the general public with these methods, supplemented by careful scientific treatises of the kind put forth by Mr. Fisher to train mathematicians in the more advanced theory of the work.

I do not mean by this that we should encourage those who do not properly understand statistical mathematics to try to use them. Such work would probably produce startling results. Rather, I would like to see the public taught that such methods exist and produce truer results, so that it will come to accept the necessity for their use as it does engineering mathematics, or the reserve in life insurance, or generally the use of mortality tables, when it really does not understand in any but the most superficial way the use that is made of them. The public is generally disposed to look upon statistics as a field where every man can and should make his own analysis and deductions from the figures as they stand and where "hifalutin'" mathematical formulae are all "tommyrot." Too many so-called "practical" statisticians are prone to encourage this idea.

I must admit that the form of theorem referred to by Mr. Fisher would have been simpler for use in preparing the table, although, perhaps, the process might not have been entirely as clear to those with less experience, for whom the paper was in part designed.

I am also disposed to agree with Mr. Fisher that we can get much further by frequency curve methods such as he uses than this paper attempts to carry us. In the hurly-burly of a rate revision as carried on at the present day, it is hopeless to undertake such work, for the frequency curves would have to be prepared practically for each of the several sets of data before the committee. It has seemed to me, however, that there was need of a simple rule of thumb which would be safe, which might act as a restraint against, on the one hand, too great confidence in insufficient data, and, on the other, too great fear of serious results where the statistical data, even of considerable volume, runs counter to preconceived notions. A rule to meet present day requirements must therefore be given in the somewhat blind fashion in which the problem has been attacked here, without precise knowledge of the

particular data to which it is to be applied. If it give a safe indication of the dependability of the data, then all the purposes for which it was undertaken have been carried out.

A large part of Mr. Fisher's discussion is not directed particularly at the subject matter of the paper, but is devoted to the use of statistical data quoted in the paper in illustraion of frequency curve work based upon the use of Charlier A and B type curves. This is very interesting and useful, and I think Mr. Fisher deserves our thanks for having given it to us. Although I have not made extensive search, I have not been able to find anywhere clearly stated the fundamental equations of the Charlier A and B type curves used by Mr. Fisher and their law of development. Might I suggest that Mr. Fisher would do a most useful service for us all if he would give us in compact form the equations of these curves and their developments?

Mr. Leslie takes issue with my suggestion in proposing that we might express our probabilities with reference to payroll exposure, and to prove that the suggestion is illogical uses a rather curious and admittedly hypothetical illustration. It would require, perhaps, no more forced illustration to prove that our present method of stating the probabilities of life and death are illogical. Of course, if we were to write insurance of the type Mr. Leslie uses on a payroll basis and attempt to express the probabilities as I have suggested, we would obtain results which are not interpretable in accordance with our established theory of probabilities. But this is merely a demonstration of the unquestioned fact that if we wish to use our present developed theory of probabilities our fundamental definitions must be adapted so that they will not be in conflict with such a theory. I am disposed to believe that throughout the general range of practical use the suggestion I have made for defining our probabilities squares with the fundamental law.

THE THEORY OF EXPERIENCE RATING—ALBERT W. WHITNEY.
VOL. IV, PAGE 274.

THE PRACTICE OF EXPERIENCE RATING—G. F. MICHELBACHER.
VOL. IV, PAGE 293.

WRITTEN DISCUSSION.

MR. W. W. GREENE:

Not so very long ago I contributed a paper to the *Proceedings* of this Society, which paper was in opposition to experience rating. My objections to experience rating were largely upon the ground that as then practiced it was contrary to the basic principles of insurance.

There is no doubt that present methods of determining manual rates are not sufficiently elastic to permit of even substantial justice to many individual risks, unless some such method as experience rating be employed. However, as between the inequity resulting from the absence of an experience rating plan and the unsound demoralized condition which follows from the use of a weak experience rating plan, I would unhesitatingly choose the former.

I have felt that the plan of experience rating developed by the Actuarial Section of the National Reference Committee on Workmen's Compensation Insurance was worthy of a trial, chiefly because the structure of the plan is, in my opinion, not inconsistent with the fundamental principles underlying workmen's compensation insurance. In taking this position I am fully mindful of the fact that in all probability we are a long way from anything approaching perfection in many important details of the plan.

Mr. Whitney's brilliant mathematical investigation, undertaken in conjunction with the work of the Actuarial Section, and for its benefit, was, I believe, the first notable attempt to place experience rating upon a basis consistent with the mathematical theory of probabilities, a step which necessarily would have to be taken before experience rating could be anything worthier than a more or less orderly method of juggling with payrolls, loss ratios and rates.

To discuss thoroughly from all important technical and mathematical angles the papers presented by Mr. Whitney and Mr. Michelbacher would require, not only considerable time and space, but also, as regards Mr. Whitney's paper, a more complete knowledge of certain branches of mathematics than I claim to possess. In fact, it was my own unfamiliarity with the tools employed by Mr. Whitney that led me to undertake an investigation of my own, which resulted in a suggestion to the Actuarial Section embodying the working formula which has been finally adopted for general practical use.

For the benefit of those who have not yet had an opportunity to read the papers in question, I may say that Mr. Whitney points out that in order to employ experience rating, it is necessary that there be such a thing as "risk experience." In other words, the particular risk must have an experience distinct from that of the class to which it is assigned. This condition does not obtain in the case of ordinary life insurance, nor in the case of fire insurance. Therefore, experience rating can be applied to workmen's compensation insurance, certain forms of liability insurance and group life insurance, only.

The problem of experience rating is defined as that of how best to weight two usually conflicting bits of evidence, namely, the indication of the class experience and the indication of the individual risk experience. In this connection it is pointed out that the credibility of the risk experience increases with the exposure (meaning number of employees) and also with the degree of hazard.

The greater part of Mr. Whitney's paper is devoted to the development of formulæ for the most probable value of the true hazard of true rate for the individual risk in view of the experience indications of the class and risk, upon the theory that risks within a class are distributed by hazard in accordance with the so-called normal frequency curve.

Mr. Michelbacher's paper is, as he says, devoted to the "development of a practical plan from fundamental theoretical principles."

Neither of these papers touch upon one important phase of the experience rating problem, namely, the manner in which the losses are to be computed for the purposes of the plan. When the Actuarial Section of the National Reference Committee was in session upon this subject, I expressed myself as of the opinion that in order to place the plan upon a sound and non-discriminatory basis, it was necessary to provide an uniform method of computing losses. Such an uniform method has in fact been adopted for use in connection with the plan in Pennsylvania and New Jersey and I believe such a step is being considered in other states.

In the early days of schedule rating, the schedule contained certain items which were called "discretionary charges" and "discretionary credits." Considerations of soundness have dictated that these discretionary portions of the schedule be eliminated. In the same way the experience rating plan of the future will, I believe, be free in so far as is humanly possible from that discretionary element arising from the valuation of losses by exercise of personal judgment only. In fact, we cannot consider the experience rating problem as solved by any means until we have, not merely a tabular method of valuing outstanding losses, but such a tabular method tested and corrected in the light of actual experience.

On the first page of his paper Mr. Whitney says: "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."

A great deal of emphasis has been placed by advocates of experience rating upon its importance as an incentive to accident prevention. I cannot see where there is any ultimate ground for conflict between the advocates of equity and the advocates of accident prevention as purposes for the experience rating plan. A good experience rating plan will encourage the prevention of accidents and will also deal more equitably with the individual risk than is possible without such a plan.

I have heard it said that the present experience rating plan will not be of any value as a means to accident prevention because the experience rate is calculated in terms of losses instead of in terms of number of accidents. Accident prevention and prompt and efficient medical service are paramount in the workmen's compensation field. If the present experience rating plan is not so adjusted as to encourage the elimination of unnecessary loss of human life and efficiency, it should be so amended as to serve these higher purposes of the business, which are of greater value even than is equity as between risks. I am inclined, however, to the view that with losses computed upon a tabular basis the experience rating plan can be used as an accident prevention incentive if an affirmative attempt is made to do so. It seems to me that there is not a vast difference between the so-called "weighted accident frequency" and the tabular method of valuing losses. I believe that such difference as exists is more a matter of terminology than anything else.

Mr. Whitney makes some very interesting comments as to future developments of fundamental theory in the rating of workmen's compensation risks. He says that "the time is now come when there should be a complete reconsideration of the manual system, the schedule system and the experience system in the effort to develop one thoroughly concatenated and consistent rating system. This involves the necessity for a thoroughgoing analysis of the logic and philosophy of rating."

Mr. Whitney has done well to emphasize the importance of regarding the rating of workmen's compensation risks as one problem. In the past the general viewpoint has frequently, in fact, usually, been sacrificed to the somewhat biased view of the specialist, whether underwriter, engineer, actuary or statistician. Admitting the desirability of an abstract study such as he suggests, I would say that it is of paramount importance to bring about a situation wherein it shall be the recognized task of some group in the compensation insurance profession to periodically review and reconcile current developments in methods of rating compensation risks, with the avowed purpose of coördinating all these developments into one consistent and logical system.

Personally I do not anticipate revolutionary progress, but rather evolutionary progress in this field. I feel that excellence in the system of rating compensation risks will, for the most part, come through the unremitting and conscientious labor of members of

this profession, and their associates in the compensation business, provided that adequate agencies exist for the purpose of coördinating the results of individual effort.

It seems to me that what we need even more than a reconsideration of fundamental theory is the establishment of a more effective clearing house for principles than the present National Reference Committee on Workmen's Compensation Insurance has proved to be. The reorganization of this body into the proposed National Council will, I trust, successfully take place, thereby marking, I believe, the beginning of a new era in the compensation business.

If the National Council becomes an effective organization we shall have in the central committee of that body, which will be termed merely the "National Reference Committee," an agency for supervising and coördinating the work of technical committees, such as the engineering committee and the actuarial committee. We shall not have, as we had during the past year, two "National Reference Committees" operating independently with no means of reconciling conflicting results. The new organization will have an executive committee in the Bureau Managers, who see to it that the time of the central committee is not taken up at the regular quarterly meetings by discussion of proposals which are not based upon the results of sufficiently thorough investigation. It should be the function of this committee of managers to anticipate problems and make sure that such problems are not lost sight of until it is a case of "locking the barn after the horse is stolen," as I fear has been too often the case during the present national emergency.

As a post-script to this discussion I am appending for the information of members of the Society copy of the memorandum which I submitted to the Actuarial Section, National Reference Committee on Workmen's Compensation Insurance upon February 12, 1918. This memorandum embodies the method by which in the first instance I arrived at the present working formula of the experience rating plan.

My efforts to develop a working formula were moved by the fact that the formulas suggested by Mr. Whitney presented serious difficulties in their practical application. In fact, the formula of Mr. Whitney, which was given most attention by the Actuarial Section, would have required the use of very extensive tables requiring a considerable volume of clerical labor for their computation, which labor would have to be repeated in case of any revision in the constants employed.

It occurred to me during the course of Mr. Whitney's investigation that if what might be termed less advanced methods were employed a simpler formula might be derived. Considerable labor finally revealed that, assuming that the risk indicated premium and the class indicated premium are both the fortuitous results of the play of a true pure premium common to both risk and class, an expression could be written in terms of simple algebraic probabili-

ties for the probability that a given value $x$ is the true pure premium for both risk and class.

The most probable value of $x$, that is, the most probable true rate for the individual risk, would upon this hypothesis be that where the above referred to expression assumes its maximum value. A determination of this maximum indicated the most probable value of $x$ to be

$$\frac{mP + np'}{m + n}$$

The above formula was at once recognized as that for the weighted average of the risk and class indicated rates. The memorandum addressed to the Actuarial Section follows:

"A BASIC FORMULA FOR EXPERIENCE RATING.

"Mr. Whitney has pointed out the usefulness of considering the experience rating problem as a case in inverse probabilities, and has developed expressions for the most probable value of $x$ (the 'true' rate of the individual risk) upon the assumptions that the individual risks in a given class are dispersed (as to true rate) in conformity to the 'normal law,' and that the true rate $x$ operates to produce $p'$ in the actual experience in accordance with the conventional algebraic theory of probabilities. Mr. Whitney has also pointed out, to a degree at least, the substantial conformity of the dispersion under the 'normal law' to that produced by the expansion of $(p + q)^n$.

"As a case in inverse probabilities, the experience rating problem may be approached from a slightly different angle in such a manner as to avoid any assumptions as to the distribution of individual hazards within the class and materially simplify the resulting formulæ, both in form and application.

"The class pure prem., $P$, and the risk indicated pure prem. $p'$ are both clues to the true risk pure prem. $x$. The most probable value of $x$ is that which implies the maximum of consistency between the risk experience and that of the class. What then is the most probable value of $x$, where $x$ represents the true prem. of both risk and class, and where $P$ and $p'$ are both the more or less random results of the operation of the true pure prem.?

"The probability that $x$ is the *true* pure prem. in both risk and class, where $P$ and $p'$ are the *indicated* pure prems. in class and risk, respectively, is

$$\frac{{}^{m}C_{mP} \cdot x^{mP}(1-x)^{m-mP} \cdot {}^{n}C_{np'} \cdot x^{np'}(1-x)^{n-np'}}{\text{Sum of numerator for all values of } x}$$

where $m$ corresponds to class exposure,

where $n$ corresponds to risk exposure.

"Then the most probable value of $x$ is that which obtains when the above fraction is at its maximum, or when

$$x^{mP+np'}(1-x)^{(m+n)-(mP+np')}$$

is at its maximum, as neither the other factors in the numerator, nor the denominator, vary with $x$.

"The required condition is satisfied if

$$(mP+np')\log(x)+\{(m+n)-(mP+np')\}\log(1-x)=\text{maximum.}$$

Equating first derivatives to zero,

$$\frac{mP+np'}{x}-\frac{(m+n)-(mP+np')}{1-x}=0$$

$$x\{(m+n)+(mP+np')-(mP+np')\}-(mP+np')=0$$

$$x(m+n)-(mP+np')=0$$

(1) $$x=\frac{mP+np'}{m+n}$$

or if

$$Z=\frac{x-P}{p'-P}$$

$$Z=\frac{\frac{mP+np'}{m+n}-P}{p'-P}=\frac{n}{m+n}$$

or, more significantly,

(2) $$Z=\frac{nP}{mP+nP}$$

"The formula (1) above is readily recognizable as the weighted average of the risk experience and that of the class. It is not, however, necessary or desirable in practice to make the weight actually assigned to the class experience depend upon the volume of payroll upon which the manual rate is based. It is better to take a standard earned prem. ($MP$) as reflecting the weight assigned to the class experience. This can readily be adjusted, and is a tangible concept with which we are all familiar. Probably even the assured can understand the logic of giving his experience more weight than that of the class where the earned prem. is over a certain figure, and less where under such figure.

"In practice it is most convenient to use formula (2)

$$z=\frac{nP}{mP+nP}$$

deriving $x$ by the formula

$$x = P + Z(p' - P). \tag{3}$$

"These formulæ can be readily applied, readily adjusted, and, I believe, readily justified regardless of whether the manual rate be split, and regardless of the number of splits, if any. Their use necessitates a minimum of tabulation."

MR. ARNE FISHER:

"Inverse Probabilities"—or, to use the more correct name, the Principle of Bayes—have for years been one of my pet subjects in the mathematical theory of probability. I therefore trust that Professor Whitney and Mr. Michelbacher will pardon me for asking a few questions for the purpose of eliminating certain doubts which have occurred to me in reading their papers in the last number of the *Proceedings*. Possibly these doubts are due to me not having understood the methods which have been employed. Certain fundamental differences remain, however, which I should like to see explained in a more detailed manner, and I must therefore ask the indulgence of the members if I will have to go to some length in explaining the nature of such differences.

I have heard several university professors and several academians lecture on the Principle of Bayes. I must confess that I have never been able to grasp what these learned gentlemen really were driving at. Being a rather stupid fellow I therefore decided to read the available literature on the subject. Right here I made a sudden discovery, and rather a startling one at that. I found that the learned savants themselves by no means agreed about the so-called "inverse probabilities." Thus Professor Chrystal, the eminent Scotch mathematician, in an address delivered before the Faculty of Actuaries advised "practical people like the actuaries to bury the laws of inverse probabilities decently out of sight, and not embalm them in text books and examination papers." The Danish astronomer and actuary of the Danish Government Life Assurance Institution, Dr. T. N. Thiele, one of the foremost authorities on the theory of statistics and observations, speaks in his "Theory of Observations" of "the fallacies underlying the Principle of Bayes and the determination of a posteriori probabilities by a purely deductive process." This was rather some interesting opinion which strengthened my budding suspicion that there was something rotten—not in the state of Denmark, but in the minds of the learned gentlemen in the lecture rooms of the universities. I therefore began to study the whole literature of the "inverse probabilities," not alone in English, but in French, Italian, German, Dutch and Scandinavian as well. One of the first things I found out was that the name "inverse probabilities" in itself is a great misnomer. Its use is limited to certain English writers—it is, for instance, not

used by the Pearsonian school—and was probably first introduced by De Morgan. The principle was first discovered by the English clergyman, Bayes, and published by his friend, the actuary, Dr. Price, as a posthumous contribution in the *Transactions of the Royal Society* for 1763. Bayes stated the rule in a rigorous manner and in a very general form. Later writers, especially English and American mathematicians, have, however, lost sight of the true Principle of Bayes and substituted in its place a false, or, to be more exact, a special case of the exact principle under what they, like Professor Whitney, call "inverse probabilities." This special case of the principle of Bayes makes use of what in logic is known as "*insufficient reason*," or what Boole aptly called "*the equal distribution of ignorance.*" Now, in the great majority of cases it is absolutely fallacious to use this principle. Yet, scores of mathematicians insist upon using this false hypothesis as a basis for their computations. In my book on "Probabilities" I have shown that on the basis of the principle of "insufficient reason" we can prove that a $x$ year old person is sure to die inside a year after we have observed that out of a group of $m$ persons, all aged $x$, *none* had died during the year. This paradoxical result arises from the fact that we have assumed on the basis of "insufficient reason" that it is equally likely that there will die 0, 1, 2, . . . or $m$ persons during the year. Many mathematicians insist upon making this rather absurd hypothesis, although they perfectly well know that it is far more probable that, for instance, 90 per cent. of a large number of, say, forty year old persons will survive one year that no one or every one will die during the year. No wonder that we encounter absurd results if we use absurd hypothesis. The principle of insufficient reason can, in my humble opinion, not be used except as a test for variation due to random sampling. If fundamentally different causes are at work in the different sample groups or complexes from which the observed event may originate, it is not permissible to use the special rule of Bayes, which is based on the above mentioned hypothesis.

Professor Whitney in the beginning of his article says:

"There would be no experience-rating problem if every risk within the class were typical of the class, for in that case the diversity in the experience would be purely adventitious. The problem arises out of the necessity of assessing the degree to which the disparity between risk experience and class experience reflects a real divergence between the true risk hazard and the average hazard of the class rather than mere chance."

It appears from this that Professor Whitney intends to investigate a fundamentally different problem than that of random sampling, and that if we were to use the principle of Bayes we should make use of the most general case where the probabilities of existence of the various complexes can not be considered as being equally likely. Yet later on Professor Whitney makes use of the special Rule of Bayes when he says:

"From this point of view one value of the real hazard of the class will be as probable as another."

It seems to me that this is a tacit assumption of the principle of insufficient reason, and that Professor Whitney bases his calculations which follow upon the same fallacious hypothesis upon which numerous other applications of the Rule of Bayes have suffered a total shipwreck.

Admitting for the present that the course followed by Professor Whitney is permissible, there remain, however, several other difficulties to be explained in the development of the various mathematical formulas in his paper. I do not pretend to be an expert judge on the subject of symbolic logic. I tried to attempt to read the massive volumes of Whitehead's "Universal Algebra," but found the task too heavy for me. Pearson in an article in *Biometrika* has taken Mr. Yule severely to task for his extended use of symbolic logic and claims that the method tends to obscure the fundamental problem in a mass of bewildering detail. There are, however, several equations, which Professor Whitney derives by ordinary algebraic methods, and on which I feel better qualified to speak than on the methods of symbolic logic.

On page 276 we find, for instance, the following statement:

"From a general knowledge of conditions we are safe in assuming that this law as a first approximation may be taken to be of the normal type. There will doubtless be some skewness. . . . The standard deviation may be taken as the measure of dispersion."

I do not wish to discourage Professor Whitney, but I feel that I can not agree with this statement. First of all, the Gaussian normal error curve which Professor Whitney insists upon using throughout his entire investigation does not even approximately represent the true distribution when the probability of the happening of an event is small. Whenever the probability of the happening of an event is less than 0.01 the Gaussian curve will surely not serve as a good representation of the distribution around the mean, unless the number of exposures approach infinity as a limiting value.

From practical experience from a large number of computations relating to various frequency distributions I know this to be true. If some of my readers are inclined to doubt my statement I can offer additional evidence by referring to the writings of Pearson and his assistants. Thus in speaking about the representation of deviations by means of a Gaussian curve, we find the following remark in the well known mathematical-statistical journal *Biometrika*, edited by Pearson:

"It is not till we get something like 30 out of 1000 in a cell that we can trust the Gaussian curve to give us at all a reasonable approach."

Again, Mr. Greenwood writing in *Biometrika* has this to say about the representation of the point binomial by a Gaussian curve:

"Its limitations are obvious. If either $p$ or $q$ be very small unless $m$ is very large indeed, and for all values of $p$ and $q$ when $m$ is small, the normal curve does not approximate closely to the binomial."

In the same article Mr. Greenwood states:

"This rule certainly applies to all cases of $m$ less than 300 or 400 and $p$ or $q$ less than 0.1."

From this it would appear that in all questions of experience rating where the probabilities of the happening of an event certainly are much less than the above values given by some of the most eminent biometricians, and where the values of the units exposed to risk also are comparatively small, it is absolutely out of the question to use the Gaussian curve.

I shall illustrate this by an actual example. Suppose that an event has happened once in 100 trials. What is the probability of its happening 0, 1, 2, 3, 4 or more than 5 times in a second sample of 50 trials?

Evaluating the results by an exact computation and also by means of the normal curve with standard deviation $\sqrt{mpq}$, we obtain the following results:

| | Exact Values. | Values from Normal Curve. |
|---|---|---|
| Chance of 0–2 successes ..... | 0.8938 | 0.9202 |
| Chance of 3–5 successes ..... | 0.1007 | 0.0022 |

This strikingly illustrates how the Gaussian curve underestimates the probabilities. The estimate of the chance of 3–5 favorable events is in this particular example 50 times too small. I leave it to my readers to decide whether such an error can be called small.

I suppose that most of the members of this society will agree with me in the statement that it is useless to deal with sample payrolls of less than 500,000 or 1,000,000 in extent. According to this limit most of the figures given by Mr. Michelbacher will seldom exceed an exposure of 100 individual samples. Moreover, in such sample sets it appears that the probabilities are very small. The pure premiums per $100 in the tables on page 301 seem on the average to be about 1.50 and go seldom over 2.57. This would give a value of $p$ or $q$ considerably less than those quoted by Pearson and Greenwood and thus completely eliminate the use of the Gaussian normal curve, as advocated by Professor Whitney.

Matters do not become better when Professor Whitney says: "The standard deviation may be taken as a measure of dispersion." It is only in a true Bernoullian series, that is to say, a series of sample sets wherein the probability for success remains constant for all sets, that this is permissible. The dispersion equals in this case the standard deviation. Professor Whitney, however, at the very beginning of his brochure admits that there are variations from one sample set to another. In fact, it is such variations that

his methods intend to discuss. We are therefore dealing with a typical Lexian Series and not with a Bernoullian Series. Now it is a well known fact, which can be proven by simple algebra, that the dispersion in all Lexian Series is greater than in the correlated Bernollian series. It often happens that the dispersion in a Lexian series is four to five times as large as the corresponding Bernoullian dispersion or standard deviation. The measure of dispersion as advocated by Professor Whitney is therefore in all cases too small. This means again that his computed values of the probability of happening of some of the greater deviations from the mean value will become too small.

Taken as a whole, it appears to me that it is much more important at the present stage of our collected statistical data to investigate the variations due to purely random sampling than try to determine basic differences by elaborate mathematical formulae. I personally am of the opinion that most of the differences which the subsequent paper of Mr. Michelbacher determines as basic or inherent deviations from the class hazard are nothing more than the results of random sampling, or if you prefer another expression, the results of pure chance. This opinion is rather strengthened by the complete omission of the computation of the important statistical quantity known as the "*mean error*" in the tables of values of $z$ in Mr. Michelbacher's table on pages 311–313. As far as I can judge this $z$ is a statistically determined quantity and is therefore afflicted with errors due to sampling. Unless the mean error is small in comparison with the value of $z$ itself, it will be of little value to attach much importance to the latter value, and I think that common prudence ought to have led Mr. Michelbacher to evaluate such mean errors.

Summarizing the above remarks my objections to Professor Whitney's paper may briefly be expressed as follows:

1. The application of the special Rule of Bayes as based upon the principle of insufficient reason can only be used to investigate variation due to random sampling and not to investigations of basic or inherent differences. For this reason it appears to the present writer that the results given in Professor Whitney's paper has reference to chance variation rather than basic differences in the various risks.

2. The assumption that the normal error curves express the frequency distribution around the mean can not be considered as even approximately true in cases where the probability is small or the number of sample sets small. This is always the case with the majority of rates in workmen's compensation.

Moreover, as I have shown in my previous discussion of Mr. Mowbray's paper, the use of the Gaussian normal curve will even for comparatively large units of exposures of payrolls produce negative losses, which of course is an absurdity. The formulas as produced by Professor Whitney are of course correct if the hypothesis

of the normal distribution is true, but they surely can not be used if the distribution around the mean value is essentially skew. Mr. Michelbacher is therefore in my opinion to be blamed for having used the Whitney formulas in a very indiscreet manner. As I showed by an actual computation from the statistical data supplied by Messrs. Mowbray and Cogswell the left end of the Gaussian curve will correspond to negative losses. Of course, it is possible that Mr. Michelbacher can explain such rather curious happenings, but I confess they puzzle me, and I am at a loss to explain the presence of the Michelbacherian negative piece of tail of the curve (representing actual gains to the underwriter from certain death and disability losses) unless the dependents of workmen indeed were willing to pay money to have their providers killed or maimed for life.

3. The standard deviation (Bernoullian dispersion) can not be used as a measure of dispersion because the frequency distribution around the mean is a Lexian frequency curve, wherein the dispersion always is greater than the corresponding Bernoullian dispersion or standard deviation. This choice of the standard deviation as a measure of dispersion leads to an underestimate of the variations arising from random sampling.

4. It would be highly advisable to have an exact computation of the "mean error" of the parameter designated by $z$ in Mr. Michelbacher's paper.

While Professor Whitney has made a highly commendable attempt to investigate the theory of experience rating by means of the principle of Bayes, it would be of great interest to attack the same problem by means of the frequency curves and the frequency correlation surfaces from the theory of homograde statistics. Stated as a purely mathematical-statistical problem the whole question of experience rating may be given in the following compact form:

"Given a series of $N$ sample sets, each set consisting of $s$ individual samples (trials or observations), to determine the equation of the frequency curve of this series."

This gives the whole problem of experience rating in a nut shell. Moreover, the theory of the frequency curves and correlation surfaces of the homograde statistical series has of late years been fully developed and adopted to practical applications, so that we ought to have no difficulty in using the theory to insurance problems, provided sufficient statistical data have been accumulated.

In conclusion I wish to emphasize—in fact, I can not emphasize it too strongly—that I have no fault to find with the purely deductive part of Professor Whitney's theory, and he deserves the thanks of the members of our society for bringing this matter to our attention. It is only with the application of the theory to concrete practical problems that I see serious obstacles.

Every mathematical theory is in the last instance founded upon certain hypotheses. In so far as we are dealing with pure mathe-

matics as apart from applied mathematics we are at complete liberty to choose any hypothesis we please as the fundamental basis for our mathematical operations. We are then always sure to reach some answer, which from the point of pure logic always will be right, provided we in our calculations have not violated the rules governing the various mathematical operations. The situation is, however, different as soon as we leave the realm of pure mathematics and turn to applied mathematics. We are then dealing with certain concrete facts or sense objects, which form the basis of our problem, and we can not indiscriminately apply any theory we please to these basic facts unless the particular hypothesis upon which this theory is founded corresponds or agrees with such facts. If the concrete sense objects or facts underlying our problem are at variance with the hypothesis of our theory it is useless to employ the mathematical formula on our observations as expressed by statistical data, because we generally will encounter absurd results of the sort I have just pointed out.

"Mathematics is," as Huxley once pointed out, "an exceedingly fine mill." But as another philosopher, the Dane, Kroman, once has said: "We can not expect to get wheat flour from the mill after we have filled the quern with oats."

I must honestly confess that in my opinion Mr. Michelbacher has filled Professor Whitney's newly knapped quern with some rather "wild oats" and the result is accordingly. We have—if actually not gotten stones—at least got some queer sort of a porridge instead of bread.

MR. GEORGE D. MOORE:

The application of experience rating to workmen's compensation insurance has always aroused considerable interest. Early proceedings of the Society will disclose many papers bearing upon the subject, so, when the Actuarial Section of the National Reference Committee was directed to develop a plan for experience rating from the ground up, the task appeared to me to be hopeless. The results, however, justified the demand and the results of their labors, as contained in the two papers presented to the Society, seem to be the last word on the subject. As a member of the committee, I followed with considerable interest the logical and mathematical developments of the plan by Mr. Whitney and must admit that I have been converted from an opponent of experience rating to one of those who sees in it the solution of one of our most vexing problems.

Before the plan was adopted a number of actual risks were rated but, as one of the members of this Society suggested, the plan, to prove sound, must not only fit actual risks but hypothetical risks of extreme nature. Following out this line of reason and using the tests which were applied to the previous plan, I submit the following with the reasons for the results disclosed.

*Tests.*

Assuming the Earned Premium in each case to be $10,000—State New York.

1. (*a*) A risk having a rate of $.59 with non-serious losses $1,000 and three deaths at $3,500 a piece, total loss ratio 115 per cent., receives a credit of 28.17 per cent.

(*b*) The same risk with the same experience if the rate is $11.51 would obtain a credit of 7.48 per cent.

2. (*a*) Another risk having a rate of $.59 with $1,000 non-serious losses and five deaths at $3,500 loss ratio 185 per cent., receives a credit of 23.62 per cent.

(*b*) The same condition with a rate of $11.51 results in a charge of 2.6 per cent.

3. (*a*) A risk with a loss cost of $4,500 without deaths gives results ranging from 18.13 per cent. credit with a rate of $.59 to 9.47 per cent. credit for a risk having a rate of $11.51.

(*b*) A risk with losses $1,000 and ten deaths at $3,500 each, total loss ratio 360 per cent., with a rate of $.59 gives a credit of 12.25 per cent., while a risk with losses of $8,000 and no deaths, loss ratio 80 per cent., gets a credit of 1.26 per cent.

At first glance, after viewing the loss ratios, one is misled by these results; however, upon further consideration and study, one is impressed with the absolute fairness of the plan, for, taking the case of 1 (*a*), where the expected cost of fatalities to all other losses is in the proportion from 1 to 4, while in 1 (*b*) the relation is about fifty fifty and considering the basic assumption that the greater cost of fatalities should be distributed evenly over all assureds and little debit or credit being given for this portion of the experience, it will be readily seen that in the first case that the proportion of the rate to meet the all other losses being 80 per cent., and in the latter 52 per cent., the credit would naturally be reduced as the results of the plan disclose. Again the small premium of the risk in question is an indication that few if any death cases are expected and it therefore follows that the three deaths must have been caused by a disaster or some other condition not necessarily inherent in the risk with a great probability of its non-occurrence in the immediate years following. The cases of 2 (*a*) and (*b*) can be analyzed in the same manner with the same result. Case 3 (*a*), however, presents another phase of the same subject. Here the proportion of premiums for all other losses ranges from 80 per cent. at the $.59 rate to 52 per cent. at the $11.51 rate, but the all other losses being much heavier, *i. e.*, $4,500 as compared with $1,000 for type 1 and 2, the credits on the same assumption used in the discussion of these latter types should be materially smaller. This is also disclosed by the application of the plan. Case 3 (*b*) is still another type. It is reasonable to expect that with an attempted even distribution of deaths over all risks and a slight debit or credit therefore that the percentage of credits will be reduced when all

other losses are increased from $1,000 to $8,000 and that they shall not be so great as is the case of 1 (*a*). The effect, however, of the heavy increase in fatality from $10,500 to $35,000 should and does have a decided effect upon the size of the credit. All of these examples tend to indicate that the method of determining whether a risk is desirable or not depends upon the loss ratio is now obsolete when the results of the present application of the experience rating plan are under observation and the effect of the application of the plan to these extreme hypothetical cases appears to give fairly good results in the light of general reasoning.

The adoption of an experience rating plan which seems to prove so satisfactory in all respects should tend to bring about the elimination of many superfluous classifications in the very near future. It is obvious, as Mr. Whitney points out, that "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." The application of the plan having rated the risk properly, it would seem that broader groups of classification could safely be determined upon without working any substantial injustice to any group of assureds. Take the case of the Wood Products, Schedule 15, Group No. 368 of the new manual grouping:

| | No. | Symbols. |
|---|---|---|
| Organ Building—pipe, including setting up at the place of delivery | 2920 | BR |
| Organ Building—cabinet or parlor | 2921 | BR |
| Musical Instrument Mfg.—wood | 2922 | BR |
| Piano Mfg. | 2923 | BR |
| Piano Action Mfg. | 2923 | BR |
| Piano Forte Case Mfg. | 2925 | BR |
| Piano Keys Mfg. | 2926 | BR |
| Piano Players Mfg. | 2927 | BJ |
| Phonograph Mfg. | 2928 | BR |
| Piano and Piano Players—assembling of manufactured parts and finishing only (N.P.D.) | 2929 | BJ |

These ten classifications could undoubtedly be merged into one and a rate determined for the group. Any variations from these rates in risks subject to experience rating could and should be corrected by the application of the plan. Also let us take the Construction, Schedule No. 26, Groups 606 and 607, Excavation and Dredging:

EXCAVATING AND DREDGING.

| | No. | Symbols. |
|---|---|---|
| Cellar Excavation (no caisson or subaqueous work), including digging holes and filling them with concrete for foundations for buildings | 6220 | DM |

| | | |
|---|---|---|
| Excavation—bridge foundations, retaining walls and bases of dams—no caisson work or cellar excavation (no blasting) | 6221 | DM |
| Pile Driving—building foundations only | 6222 | DM |
| Dredging—N.O.C. (all operations except rock drilling and blasting) | 6223 | DB |
| Dredging—excavation by means of suction dredges only. All operations except rock drilling and blasting | 6224 | DG |
| Ditch Digging—no sewer or canal building or excavation for water or gas mains (no blasting) | 6225 | CR |
| Cellar Excavation (no caisson or subaqueous work and no blasting) maximum depth of excavation 12 ft. (N.P.D.) | 6227 | DA |
| Canal Construction (excluding barge or ship canal construction)—all operations in connection therewith, except railroad operations, bridge building, caisson work and wrecking. | 6361 | DJ |
| The above excepted classifications to take full manual rates. For contracts involving the performance of dredging work only—no rock work, no blasting and no other operations of any nature whatsoever—apply the rate for the classification "Dredging, by floating dredges." | | |
| Canal Construction—barge or ship | 6363 | (a) |

Group 607.

PILE DRIVING, DAMS AND DOCKS.

| | | |
|---|---|---|
| Dam Construction—excluding the construction of concrete dams (no blasting) | 6002 | DM |
| Pile Driving—including timber wharf building | 6003 | DM |
| Marine Railway—Construction | 6004 | DM |
| Jetty and Breakwater Construction (no blasting) | 6005 | DM |
| Dry Docks—construction (no blasting) | 6008 | (a) |
| Waterworks—construction of pumping stations, dams and reservoirs | 6010 | (a) |

Of the fifteen classifications in the above groups, six could be safely merged. The condensing of the classifications would be of immeasurable value in the compilation of Schedule Z, which has now become such a necessary part of rate-making, and, as experience rating has apparently become a permanent fixture in compensation insurance, why not extend its use in this manner, thus conserving considerable labor and expense which could be more profitably applied to the analysis of individual accidents so sorely needed at this time.

One of the most radical changes in the method proposed in the plan was the application of experience rating to the policy premiums as effected by schedule rating, the argument being advanced that experience rating, of necessity, discloses results of physical conditions, morale and every other loss-producing factor and that it was therefore inequitable and gave far too much weight to debits and credits when both plans were applied separately. After careful consideration of the matter, however, I have come to the following conclusion: As experience rating now applied under the new plan may nullify the effect of schedule rating and for the purpose of simplifying the work of experience rating, I am inclined to believe that the experience rate should be applied to the full manual premium, after which the debit or credit, depending upon the change in the effect of the application of schedule rating as disclosed at the date of last renewal and that disclosed at the date of current renewal, should be applied to the experience rate determined above. The application of schedule rating in its full detail should, of course, be applied to non-experience rated risks.

MR. WILLIAM J. GRAHAM:

"The Theory of Experience Rating" by Albert W. Whitney and "The Practice of Experience Rating" by G. F. Michelbacher admirably complement each other and ably present the subject of experience rating. The Society is fortunate in having this wealth of authoritative information on experience rating to discuss at this meeting. The subject, however, is so broad that even the two papers read together must naturally presume knowledge on many important references not included in the text of the papers, such, for example, as the investigation of the theory of experience rating and the theory recently adopted by the Actuarial Section of the National Reference Committee on Workmen's Compensation Insurance.

In the assumption that experience rating is now established and destined to be a factor in life underwriting when there is a risk experience to be considered as well as a class experience, the authors of the papers have avoided the tedium of repeating defenses in behalf of experience rating, which simplifies their subject and enables them to go directly to the point of the application of the theory. While these papers are differently titled, both deal with theories, and both in a measure deal with practices. The practical references to underwriting problems which have introduced the need of some such refinement in fixing premium costs as is found in experience rating, constitute a clean-cut analysis of practical problems as they are today in the various branches of underwriting in which there is a risk-experience as distinguished from a class-experience.

Experience rating has arisen out of the necessity for a greater equity in fixing premium rates in workmen's compensation insur-

ance, where manual rates and even schedule rating have been more or less qualified failures. Viewing manual rates purely as class rates, and schedule rates as a refinement of the general class rate, to those more intimate studies with reference to some factors of risk affecting the class hazards, there remains to be evolved a system that would take into account risk-experience where the same could be done with profit and justice because of the character of the hazard and the sufficiency of the exposure within the particular risk. All forms of underwriting must remain a class-underwriting one to the extent of dependency upon averages. Carried to its highest refinement, experience rating must be restricted to the field in which a balance may be properly obtained between the risk effect and the class effect, so that while taking into account in a proper degree the risk affecting the same must ever be balanced and subordinated to the class effect. The four elements of balance are exposure, hazard, degree of concentration, and credibility of the manual rate.

It is quite conceivable that a particular risk may reach proportions through frequency of hazard and quantity of exposure to the point where it constitutes a complete class. If this risk is homogeneous and eliminates all the problems which must otherwise arise with reference to balancing class effect and risk effect, it will in itself establish the perfect balance and accord between the two.

While experience rating offers an incentive to insurers to keep down losses within the establishment, it does this in the indirect way of having such saving reflected in the new rating. The schedule rating system, however, directly credits methods adopted for saving life and limb in a more feasible way, and thus contributes an element of value which we must conserve so far as compatible with the broader principles of life underwriting because of its unquestioned value in the prevention of accidents.

One way of viewing and applying the theories expressed in these two papers would be to consider the manual rate as the basic, unrefined class-rate; to view the action of schedule rating on the manual rate as a refined class-rate. This rate is adopted as the basic rate and subjected to an intelligent experience rating plan which would supplement a manual rate supposedly refined to the greatest practical point by balancing this rate with an experience weighted for the particular risk, so far as the incidence of exposure or hazard warrant. I would point out here, however, that the result of this action merely produces a new rate which in itself may be viewed as a prospective rate. There is nothing in the theory or practice of experience rating as announced that would correct any mistakes of the past as developed from the actual experience. This experience is merely to be utilized to quote a new rate, which rate in turn will be subjected one year hence to the same type of mathematical analysis with reference not only to adjusting the year's experience by a system of debits or credits, but to further weigh in the

risk experience in fixing a new rate for the class. In theory this system would mean a large accuracy in equitable adjustments, if it is carried on through the years with factors which permit of increasing risk weight as against the decreasing class weight. None the less, practical conditions concerning changes in industries and in the methods of manufacturing, besides voluntary changes from one type or company of underwriters to another, all introduces difficulties which emphasize the point that the experience rating formula is not, properly speaking, a retrospective formula except as it is compared with a formula for class rating such as manual rates or the more refined schedule rates. It is possible to apply the experience rating formula in thoroughly retrospective fashion to give weight to the risk experience in a series of debits by calling for the payment of additional premiums at the end of the year on the one hand, and on the other allowing credits with reference to the particular year's experience. In practice, however, this theory would have many disadvantages, not the least of which would be the great difficulty of attempting to collect greater debits than may be deemed to be due by reason of the past year's experience.

A practical way in which a retrospective experience rating can be effected is by charging a premium appreciably higher than would be called for by the class experience and permitting premium refunds at the end of the year to be effected by an experience rating formula. This formula should provide for credits to be made in excess of the average class dividend up to the full amount of the gross premium, and of appropriate counterbalancing debits. This would introduce the participating and mutual idea of underwriting into the workmen's compensation field. In itself this would be no particular departure, since there are now types of organizations attempting to apply the mutual principle, without, however, going to the point of advocating a thoroughly retrospective experience rating method of adjusting the risk-experience to the class-experience through the medium of a general increase to the class rates, to be offset by appropriate experience rating premium refunds to the end of the policy period. This general principle is announced in harmony with the general theories of the papers that experience rating is restricted to contemplate and adjust independent occurrences of a simple contingency.

Boiling down the theory of experience rating to the point of conserving the equity of the individual risk by striking a balance between the weight to be given to class experience and the weight to be given to risk experience, there remains two ways in which the theory may be applied. First, with reference to non-participating rates such as evolved and applied by Messrs. Whitney and Michelbacher and which, with apologies to that portion of Mr. Whitney's paper which refers to the theories of experience rating as a retrospective plan, I would describe it as a prospective method of rate readjustment. Secondly, the other way of applying the same prin-

ciple is that of the participating plan of not attempting the same degree of nicety with reference to the original class rate, but to increase that rate to a point which would permit of a much greater ultimate nicety of adjustment to the individual risk by the debit and credit method of premium refunds to be decreased or increased above the average refund of the class within the limits of the plan according to actual experience. An expansion of this thought would lead to the conclusion that the larger the premium the less the effect and consequently the larger ultimate justice that could be done to the class. Theoretically this is true and in practice could be true to a large degree, in fact up to the point where a variation in risk-experience would reach its maximum and extend beyond to extra hazardous classes obviously requiring additional classifications.

No one in the field of practical underwriting is looking for perfection, or has any illusion or hope as to realizing it. The business of insurance is to average up risks. In this averaging lies the security of the credits as against the premium payments of the individual who may or may not realize at all upon his premium payments, or who may or may not realize in anything like the same measure of returns for premiums paid. Professor Whitney admirably restates this principle when he says at page 281: "The fundamental theory of insurance involves this, that, at the point when the effort to analyze and differentiate the hazard of various risks has been carried as far as is deemed feasible, the risk in each residuum shall be treated as of equal hazard. This means therefore that each risk shall take the average hazard of the group."

In all our efforts for risk refinement we must keep the inherent class factors in mind as a basis on which insurance endures. The papers of both Professor Whitney and Mr. Michelbacher have pointed out the limitations of the experience rating theory. While pointing out that the problem of experience rating is peculiar to workmen's compensation insurance chiefly and not found in life insurance, it makes exceptions of the potential possibilities of the problem in group insurance. Naturally in life insurance there is no risk experience since a person dies but once. It is interesting to note that the participating plan in life insurance has developed a contribution formula by which a refinement for classes as against individuals has been made possible by the application of analogous principles. The contribution formula subdivides and returns to classes of life insurance risks a premium refund in the form of annual dividends, adjusted with reference to the experience of the class as regards the different factors which go to make up the so-called dividend earnings. In group insurance, however, the risk-experience is introduced directly and the life insurance as a factor would be clearly distinguished from class-experience. I have treated this problem at some length in a written discussion of the joint paper submitted to the Actuarial Society of America, the able contribution of Messrs. E. E. Cammack and E. B. Morris. The paper

was presented to that Society in May, 1918, and is entitled "Joint Mortality Experience of the Aetna Life and Travelers Insurance Companies on Group Policies." (*T. A. S. A.*, XIX, 29.)

The problem in group insurance furnishes a particularly happy application of the theory so well developed and worked out by Messrs. Whitney and Michelbacher, since in group insurance the law is of that simple character which lends itself most readily to the application of the formulas developed. The amount of the risk is apparently not large in group insurance, as it is subject to the following regulations:

*Maximum Amount to an Individual.*—The maximum amount of insurance as to any individual in any class shall not exceed two and one-half times the average of the group (to the next $500) determined by excluding such class, but in no case shall the insurance on any individual exceed $5,000. The usual minimum amount of insurance on any individual is $500.

This again excludes the element of catastrophe, which for reasons so well developed in the papers under discussion are not properly the subject of experience rating discussions.

The Equitable Life Assurance Society issues all its group insurance on the mutual plan at participating rates, and applies to all its business at this time an experience rating dividend formula. This formula contains many of the arbitrary factors which both Professor Whitney and Mr. Michelbacher found necessary to introduce with reference to workability, and with essential reference to preserving a balance between risk-experience and class-experience which would not unduly penalize the small risk having a bad experience, which so far as its individual classification was concerned, might be considered more or less adventitious. The Equitable takes into account the size of the group, the number of years experience in each group and the constantly changing character of industrial establishments by carrying its debits and credits algebraically, checking back over the entire experience of the group except as the size of the group may increase.

Group insurance is governed by the laws of the various States of the Union applicable to individual life insurance. Thus the matters of fixing premium rates with reference to individual ages and the maintenance of individual reserves are matters of statute, and so long as they remain matters of statute would not permit of the experience-rating methods of rate making. Where the individual rates, however, are in conformity with the state standards, the theory and practice of rate making as outlined by Professor Whitney and Mr. Michelbacher may, with modifications, be used with reference to computing premium refunds or dividends. But the problem of premium refunds in a group life risk, in retrospection, eliminates many of the difficulties with which Professor Whitney and Mr. Michelbacher have made their analysis to so large an extent in the broader and more comprehensive formulas developed

by them, with workmen's compensation hazards particularly in mind.

A. W. WHITNEY:

(AUTHOR'S REVIEW OF DISCUSSION.)

Mr. Greene refers to the manner in which the losses are to be computed in the application of the experience rating plan. The computation of losses is doubtless the weakest part of the system and every effort should be made to standardize and regularize the procedure.

Experience rating has come to stay and as a method for making right rates will be more important than schedule rating. There seems to be no other field in which the necessity of making essentially right rates for the individual risk is so imperative as in the field of workmen's compensation. This new actuarial method must therefore be given as dignified and useful a place as possible.

I am very glad that Mr. Greene has added as a matter of record the memorandum that he presented to the Actuarial Committee. It was Mr. Greene's assumption which brought to light the form for $z$ which, on account of its great simplicity, was adopted for practical use. As a matter of record also I might state that the original working out of the problem was made on the assumption that the class-experience was so large as to make the class-rate entirely dependable. In the paper which was presented to the Society and which is now under discussion the treatment was generalized so as to include both points of view as special cases.

While the assumption which Mr. Greene made was valuable in bringing immediately to light a very simple formula I cannot consider that the assumption itself really reflected the facts. It is certainly not true that experience-rating as a balance between the credibility of risk-experience and class-experience depends solely or even primarily upon the relative amount of experience for class and risk. Mr. Greene assumes that all risks of the same class have the same hazard.* This of course is flatly at variance with the facts; the need for experience-rating rests upon the exact contrary, namely, that the hazard of the risk in general is not the same as the hazard of the class.

Incidentally notice should be taken of the fact that the symbols used in Mr. Greene's memorandum were adapted to those used in the earlier devlopment and do not wholly agree with the symbols used in the paper which is the subject of this discussion.

There are three distinct processes in the theory of probabilities: first, an analysis of the logical relationships among the events in question; second, the superposition upon this logical structure of

* Formula (22) reduces to the form $Z = \frac{Pn}{Pn + Pm}$ by letting $H^2 = \infty$ which is the condition that the hazard of all risks in the class is the same.

a system of quantitative relationships, and third, the making of explicit judgments regarding the magnitude or relative magnitude of the quantities involved. The first process follows out of the fundamental laws of thought; the second, out of similar quantitative laws; the third process is entirely different for it involves specific judgments regarding the nature and detail of reality itself. All three processes involve judgments but in the first two cases the judgments are so fundamental, so general, so well-established, that they are non-controversial. The structure of inverse probabilities for instance is something that is not open to question.

The main difficulty is concentrated in the third process of making judgments as indeed is the case in the more ordinary affairs of life. People on the whole reason logically, and live consistently. Their failure to meet successfully the pragmatic test of reality lies mainly in their inability to make sound judgments on which to base their reasoning or their living.

The practical value of a determination of the probability of an event is wholly conditioned by these initial judgments. The result carries with it all the qualities, the imperfections as well as the perfections, the inadequacy as well as the adequacy, of the assumptions. Here, as elsewhere in life, one of the chief difficulties lies in the fact that we forget that our results are qualified,—we take them as absolute. In the application of the theory of probabilities and notably in the application of Bayes' rule these limitations are often overlooked and the results are given a credence that is unwarranted.

So far as mere logic is concerned any self-consistent assumptions are admissible. Whether these will lead, however, to results that have any practical value depends entirely upon the question of whether such assumptions are a correct description of reality.

Among the various probabilities that follow from various assumptions is there such a thing as *the* probability par excellence? Yes, it is that probability that is based upon the best possible judgments, —that reflects most thoroughly the known facts. But who shall be the judge?

In the case of hypothetical problems where the body of assumed facts is explicitly given, and particularly where the domain of ignorance is posted with signs marked "at random" the determination of *the* probability is not difficult. In real life the situation is quite otherwise; here it is almost impossible to marshall in review all the knowledge, much of it scattering, which goes to make up the logical background. And yet with a characteristic mental carelessness we do even in such cases use the term *the* probability. Similarly in the more ordinary affairs of life we neglect in general to recognize the fact that our conclusions have strict validity only for one particular substratum of experience. In some parts of the theory of probabilities, notably in the case of inverse probabilities, this procedure is too crude. Distinctions must be explicitly real-

ized between the results that arise out of different experiential backgrounds.

There is such an air of mystery about the theory of probabilities; —it is so startling to be able to get dependable results where there is ignorance,—that it is not strange that there should seem to be some alchemy by which ignorance itself is transmuted into knowledge. Here, however, as everywhere else the fact is that what we do we do in spite of our ignorance and not because of it.

The validity of our conclusions with regard to the probability of an event is an exact measure of the extent of our knowledge. There must to be sure be ignorance if there is to be probability (rather than certainty) but this ignorance is in the nature of lacunae in a body of knowledge and it is out of this knowledge rather than out of this ignorance that the theory of probability builds a bridge across the gap.

It is true that a determination of probability based on perfect ignorance would have no value whatsoever. The attempts of a being who dropped into this world from a totally different sphere of existence to apply either certain inference or probable inference (even though he were familiar with the structural nature of logic and of the theory of probability) would be perfectly futile because of the fact that he had no underlying basis of experience out of which to construct judgments. Such a case in actual life is impossible. However tenuous may be the threads of knowledge there is nevertheless something that can be woven into a probability determination.

There is, however, in real life not only no case of perfect ignorance but there are all gradations of knowledge. For instance, here are 501 of them: An urn contains 1,000 balls which are known to be no other than black or white.

1. A ball is drawn; what is the probability that it is white?

2. 2 balls are drawn and found to be one black and one white, and replaced; a ball is now drawn. What is the probability that it is white?

3. 4 balls are drawn and found to be two black and two white and replaced; a ball is now drawn. What is the probability that it is white?

. . . . . .

500. 998 balls are drawn and found to be 499 black and 499 white and replaced; a ball is now drawn. What is the probability that it is white?

501. 1000 balls are drawn and found to be 500 black and 500 white and replaced; a ball is now drawn. What is the probability that it is white?

In the first case the only knowledge is an underlying acquaintance with the world in general; this throws a few rays of light upon the conditions of the problem. In each succeeding case there is slightly greater knowledge, until in the last case there is com-

plete knowledge with regard to the contents of the urn. The application of the theory of probabilities in each case gives the same result, namely, that the drawing of a white ball and the drawing of a black ball are equally likely and yet the credence to be attributed to the result in the different cases is very different. In each of these cases except the last a probability determination requires the making of arbitrary a priori assumptions.

The difference produced by variations in the amount of knowledge available can also be shown by a modification of the first problem.

Problem I. An urn contains 1,000 balls which are known to be no other than black or white. A ball is drawn at random; what is the probability that it is white?

Problem I–A. There are 1,001 urns each containing 1,000 balls. In the first urn all are white, in the second all but one are white . . . in the 1,000th 999 are white, in the 1,001st all are black. An urn is selected at random and then a ball is selected at random from this urn. What is the probability that it is white?

The mathematics and the result in the case of these two problems are exactly the same. In the first case assumptions are made that are the equivalent of facts in Problem I–A. It is evident that there is a larger amount of knowledge in Problem I–A than in Problem I and the result to just that extent is more worthy of credence. Problem 501 differs from Problem I–A in the addition of still further knowledge; in fact one element of probability has given way to certainty.

At the bottom of every probability determination in real life there are assumptions that are more or less arbitrary. Furthermore it is often if not usually impossible to determine just what assumptions arise most consistently from the logical background.

Both of these conditions are matched in the case of certain inference. Before experience is accumulated arbitrary tentative judgments must be made; the results do not command a high degree of credence and must be tested pragmatically. Furthermore it is often exceedingly difficult to pick out those assumptions which are most consistent with the logical background of experience.

The conclusion therefore is that the credence which is to be given to a probability determination varies with the extent of the logical background and the degree to which it is possible to express this with certainty in the form of assumptions.

Mr. Fisher's example in which 0, 1, 2, $\cdots$ $m$ deaths are assumed as all equally probable and from this assumption is derived an improbable result illustrates only what seems to be self-evident, namely, that consistency must be judged within a uniform field of experience. It is not legitimate to apply a different criterion of reasonableness to the final result from that which is applied to the assumptions. If a body of experience is called into evidence for the purpose of pronouncing the fact unreasonable that an $x$-year-

old person is sure to die during the year then this same body of experience must be called into use in passing upon the reasonableness of the assumption that 0, 1, 2, $\cdots$ $m$ persons will die during the year. In other words anyone who is qualified to pronounce the result unreasonable is also qualified to pronounce the assumption unreasonable. Similarly the result is not unreasonable if the same body of ignorance is assumed that is assumed in testing the assumptions.

It is just as fundamental in a logical problem that the same body of experience should be assumed throughout as it is that the same universe of discourse should be preserved throughout. It is not allowable to swap the horse carrying our load of experience in the middle of the logical stream.

Now to come to the particular case in point. Is it allowable to assume that one value of the real hazard of a particular class is as probable as another provided the experience of the class is set aside. This is by no means a case of complete ignorance. We have not only a fund of general knowledge but with regard to this particular class we doubtless should be able from a technical point of view even barring experience to form some idea of the probable hazard and to assert that in reality all values of the hazard were not equally probable. It would be very difficult, however, to establish an a priori law of probable hazard that would produce a consensus of approval. Furthermore, for the range of values in which our chief interest lies it is pretty evident that such a law expressed as a curve would be nearly flat.

Under the circumstances, with the additional motive of producing the greatest possible simplicity, we are certainly justified in going back to the more naïve view that all values of the hazard are equally probable. The difference in the results produced by using this law and by using some more complicated relationship in practice can certainly not be great. I doubt if it would be desirable to make a different assumption, considering the purpose of the investigation, even if more explicit estimates were possible.

It is pertinent to ask what is that purpose? The answer is that the investigation was undertaken to discover a form, a structure, for a plan of experience rating that would be in general agreement with underlying conditions and requirements. The best assumptions that were practical were to be made use of but the result itself was to be judged by pragmatic tests. In this connection I may note that this plan is being used to rate twenty million dollars worth of business and is apparently giving satisfaction.

I agree with Mr. Fisher that normal curves do not exactly describe the conditions of the problem; they were not supposed to. For the purpose in hand, however, they were exactly the instrument that was needed. This was not a theoretical investigation; it was governed by eminently practical considerations. The normal probability curve with its ease of handling marked the exact limit

to which it was feasible to go in the direction of nicety of treatment.

Even aside from this consideration, however, Mr. Fisher's criticism is not to the point. The first factor in formula (2), viz., $H'/\sqrt{\pi}\cdot e^{-H'^2(P-X)^2}$ does not affect the form (at least only remotely); it affects only the constant

$$\frac{H^2H'^2}{H^2+H'^2}=J^2;$$

but this constant $J^2$ in practice is determined arbitrarily. The second factor $H/\sqrt{\pi}\cdot e^{-H^2(x-X)^2}$ represents a frequency distribution and is not used as an approximation to a binominal form. The question of skewness for the purpose in hand is not important. The third factor ${}_nC_{pn}(1-x)^{qn}x^{pn}$ is exact, not an approximation.

I hope that somebody with time and patience will investigate the problem of experience rating from a more comprehensive point of view. I conceive that such an investigation would be not only theoretically interesting but might throw some light on the practical problem. The present investigation was inspired by the necessity of producing immediately practical results.

The term "inverse probabilities" has no great merit, neither is it pernicious. The theory of probabilities in form is reversible. In its applications, however, there are considerable differences. Usually we argue from cause to effect; in some cases, and particularly in the application of Bayes' rule we argue from effect to cause. This inversion of the more usual process was evidently what led to the term "inverse probabilities," as well as to the term, the probability of causes.

Mr. Fisher's remarks about dispersion are hardly to the point in view of the fact that no use was made of the idea of dispersion in the technical sense.

I am very glad that Mr. Moore has taken the trouble to enrich the discussion of experience rating with the submission and analysis of critical cases. What the experience rating plan ought to do is to reproduce in a systematic manner the judgment of the underwriter. These cases seem to indicate that the plan is successful in doing this.

The difficulty in using experience rating to displace complexity in classification lies in the difficulty of dealing with the small risk. The rating of the small risk, whatever its experience, will have to be kept pretty closely to manual and if the classification is in error experience rating will not go far in the way of correction.

## THE INDUSTRIAL COMPENSATION RATING SCHEDULE, 1918—E. H. DOWNEY.

VOL. IV, PAGE 325.

WRITTEN DISCUSSION.

MR. G. F. MICHELBACHER:

Everyone who has come in contact with the latest edition of the Industrial Compensation Rating Schedule will admit its superiority over the schedules which had their existence prior to 1918 and out of which the present schedule has developed. I can find no basis for criticism, therefore, of Mr. Downey's analysis of the faults of the first workmen's compensation schedules, nor have I any comment to offer with reference to the efforts of the National Reference Committee on Schedule Rating to overcome these faults and to produce a truly practicable and equitable plan of schedule rating.

I am not intensely interested in the details of schedule rating because I have had no training in the technique of safety engineering. Personally, I am content to leave these matters in the hands of safety engineers. It is their business to know what conditions are of sufficient importance to warrant consideration in a schedule and they are better qualified than I to speak both in terms of the relative values of these hazards and the methods which should be employed in their elimination. The right sort of statistical information should be available but, after all, the limitations of our experience data are such and it is so obviously impossible to obtain the exposure to individual hazards with any degree of accuracy that the actual valuation of the items of any schedule must remain largely a matter requiring the expert judgment of engineers.

I do feel, however, that the structure of schedule rating and its place in the general plan for rating workmen's compensation risks are matters in which actuaries should have a decided interest. In this connection I have two points upon which I feel at liberty to comment.

In the first place, I believe the present schedule to be too cumbersome and complicated. Efforts have been made periodically to simplify schedule rating. The injection of new ideas and the desire on the part of those participating in these conferences to experiment with new principles has, however, led to the gradual building up of a most formidable array of items and formulæ. I have no facts before me to substantiate my case, but I am certain that if it were possible to make the right sort of examination it would be discovered that there are many of these items which have

little effect on the aggregate results obtained upon all the risks which are schedule rated. It appeals to me that there are too many refinements; there are a number of items representing hazards which are not of sufficient importance to warrant the trouble and cost of attempting to measure them. In this respect the schedule is top-heavy. There must be a possible middle ground upon which to build a schedule which will efficiently and equitably measure physical hazards without the necessity of considering minute details and hazards which, as regards the broad spread of industry, are comparatively insignificant. It may be that our attempt to make the schedule universally applicable to manufacturing risks is responsible for this condition. It is possible that the fact that one schedule is applied to manufacturing risks of every description requires the inclusion of items which are important in certain classifications and of absolutely no importance in others. Nevertheless, the recent development of formulæ for the purpose of better fitting the schedule to individual classifications should remove this difficulty. I look forward to the establishment some day of a schedule which as regards industry will produce identically the same results as our present schedule but which will contain only a few items of general application to all risks.

The second point upon which I should like to speak has to do with the place of schedule rating in the general plan for rating workmen's compensation risks. At the time the latest experience rating plan was devised, the theory was advanced that the schedule may be considered as a method of refining the classification of a risk. If this theory is applied in actual practice, it results in the use of the manual first, the schedule second, and the experience rating plan third. Each of these parts of the system has a definite mission to perform.

It is the purpose of the manual to furnish a convenient method of breaking up industry into a number of suitable parts for classification purposes and to provide an average or base rate for each division. The first step in the underwriting of an individual risk involves the consultation of the manual. The classification which most nearly describes the operations conducted by the assured is located and the rate applicable thereto is taken as the basis for rating the risk. It is recognized, however, that this is an average rate—that it is based upon experience into which has been thrown individual risk experiences of all kinds ranging from the experience of superior risks to that of risks which because of physical and moral conditions are extra hazardous within the manual classification. Equity and good underwriting judgment dictate the necessity of making some variation from this average or base rate in order that the rate for the individual risk may more intimately measure the hazards and thus insure a more accurate contribution from the assured. The next step, therefore, is the application of the schedule.

The schedule measures physical conditions and according to the theory that it may be used to refine the classification of a risk, its application is equivalent to the breaking up of a manual classification. It is as if all the risks which are thrown into a heap within a manual classification were sorted into a number of smaller heaps. The schedule is supposed to accomplish this result by measuring the extent to which the physical characteristics of the individual risk vary from those assumed to be represented by the manual or base rate. After the schedule has been applied the risk is more accurately classified. It is thrown into immediate relationship with risks of substantially similar hazards where the physical conditions are comparable and it no longer takes its place at random with the good and bad risks of the classification.

The last step is then taken. The loss history of the risk is reviewed and the experience obtained in this manner is compared with the experience for the classification. There is a process of weighing the relative evidential value of the risk experience and the schedule rate and an adjustment is made which, if the risk experience is better than the average, results in a credit, or if the risk experience should indicate a contrary deviation, a debit.

This theory of rating is being followed at the present time. The manual rate is first modified by schedule. The rate thus obtained is then subjected to further modification by experience rating.

This is not the only theory which has been advanced, however, and personally I am not convinced that it is the best that can be found. There is a distinct movement in favor of reversing the order of application of schedule and experience rating. Experience rating would thus be employed as the method of refining the classification if a risk and schedule rating would constitute a means of stimulating accident prevention and of anticipating the effect of safety work.

All the hazards of a risk, both physical and moral, are reflected in the experience of the risk. Why is it necessary, therefore, to look further for a dependable criterion to use in bringing together those risks within a manual classification which represent approximately the same hazards? To be sure, the present experience rating plan does not consider the experience of all risks to have the same evidential value. Nevertheless, some suitable modification might be made if the plan were used exclusively for classification purposes—*e. g.*, greater weight might be given to the "all other" experience, and less weight to the "D. and P.T.D." elements.

Schedule rating in the past has been designed to accomplish two purposes: first, it has been used as a method of measuring physical hazards or of determining how much each physical hazard present in a risk contributes to the total loss cost; second, it has been used to stimulate accident prevention. By forcefully bringing to the attention of assured the importance of serious hazards it has more than any other agency advanced the safety first movement.

As a plan for the measurement of hazards, the schedule has not produced good results; as a booster for safety, it has accomplished wonders. Of the two purposes of schedule rating the commercialization of accident prevention is by far the more important. It is so important that I should not like to see it go. I believe that this result may be accomplished, however, without an elaborate schedule and without permitting the schedule to have too much influence as a method of measuring the hazards of a risk.

I would accomplish this by placing the schedule third in the scheme and using it merely for the purpose of anticipating the value of safety work. If we assume that the effect of guarding a saw will some day be reflected in the experience of the risk, why not permit experience rating to measure it? The difficulty is that if an employer has to wait until the experience of his risk reflects his accident prevention work he may not be particularly interested in making safety installations. The stimulus which the schedule in its present form applies to accident prevention would be lost if schedule rating were entirely discontinued. If, however, the employer were afforded the opportunity of earning a credit for a safety appliance, say, for two years following its installation, the incentive would still be there. A rule might be established, therefore, which would limit the application of the schedule to new installations and those which have been in effect one year at the time of inspection. After that the schedule would not recognize the condition and the experience rating plan would reflect the credit which the safety installation is entitled to receive. This would be measured in an actual saving in loss cost and not in some arbitrary manner based upon the assumed value of the hazard which has been eliminated or the cost of making the safety installation.

This plan has certain advantages over the one in present use. In the first place its adoption would render unnecessary a change in the experience rating modification each time the schedule modification is changed. It is customary to inspect risks as often as three times a year, depending upon the size of the premium. The present plan requires a recomputation of the experience modification as the result of each of these inspections, for presumably the condition of the risk will be found to be different else there could be no argument in favor of so many inspections. In accordance with present theory a change in the schedule modification is equivalent to a change in the classification of a risk. There is not much logic in the contention that this can happen several times a year. It is far more logical to assume that the classification of a risk is temporarily fixed for at least a term of twelve months. The character of the risk may differ from time to time because of an actively conducted safety campaign. Under the proposed plan such changes would be measured by the successive inspections and the rate modifications based thereon.

Then, again, the proposed plan would avoid the possibility of

a duplication of debits or credits. There is this danger if the schedule is applied before the experience rating plan comes into play. Let us assume, for example, that there is a superior condition in a risk. If this condition is treated by the schedule it will receive consideration when the schedule is applied. But it also will be reflected in the experience of the risk and it will thus have an influence in determining the experience rate. To be sure this influence may not be such as will double the schedule credit—it will be felt, however, and in the smaller risks items of this character may have an appreciable effect upon the adjusted rate. In the larger risks, of course, the experience rating plan will tend to reproduce the risk experience and here the schedule will have little influence. Under the proposed scheme the jurisdiction of the schedule and the experience rating plans could be clearly defined. I am not at all certain that the experience of a risk accurately reflects the effect of safety work within two years after installations are made. This is a matter which could be studied statistically and also one concerning which the judgment of engineers would be valuable. It should not be difficult to produce a rule which would result in carrying the value of safety work by the schedule until such time as the experience is capable of taking over the burden. Even though the rule were not accurate the present situation would be much improved.

Finally, the proposed plan, if adopted, would remove certain difficulties which are inherent in the present plan. At present if a subsequent inspection produces a greater credit than the original inspection the schedule ratable classifications are given lower rates. However, the fact that the average rate and the indicated risk rate come closer together produces a smaller experience modification and when this is applied to the non-schedule ratable classifications the result is an increase in these rates. It is true that the average rate for the risk decreases but the practical result is a decrease on one set of classifications and an increase on another, and this is by no means a simple matter to justify to an assured. Another condition which has confronted the Bureaus is this: under the present plan a risk cannot be completely rated until the schedule modification is available. Furthermore, all the operations of an assured within a given jurisdiction must be taken together for experience rating purposes. If there are three manufacturing plants under the control of a single assured three inspections must be made and the experience for the entire risk must be compiled before any of the constituent parts can be completely rated. If there is a controversy concerning the application of the schedule to one part of the risk, the rating of the entire risk is delayed. It is a case where the tail may wag the dog. To be sure these difficulties can be overcome but any rule which would eliminate them under the present plan would be awkward. Under the proposed plan there would be no conditions of this character, for with the schedule as the last

step in the rating procedure, any delay in determining the schedule modification or any adjustment in rates made necessary by reinspection could readily be taken into consideration.

This subject is at present under investigation by the National Reference Committee on Workmen's Compensation Insurance. I trust, therefore, that this plan and others which may be proposed will be thoroughly discussed here, as the coördination of the various elements of our scheme of rating is undoubtedly the next important problem in the field of workmen's compensation insurance which must be solved.

MR. W. H. BURHOP:

In view of the fact that Dr. Downey's paper on the above subject presents mainly a comparison of the Universal Analytic Schedule with the Industrial Compensation Rating Schedule, and offers a very complete explanation of the latter, little opportunity for a discussion is presented. The paper should give a better understanding of the subject to those not immediately connected with the preparation or application of rating plans.

Dr. Downey places a great deal of emphasis on the balancing feature of rating schedules. It is difficult to understand how we could ever have a plan that would be in balance and remain so without frequent revision of the credit and charge items. Our pure premiums, the basis of our rates, represent average conditions; so the rating schedule is intended to produce the manual rate for the average plant, and a higher or lower than manual rate for plants presenting a poorer or better than average condition. Safeguarding of machinery is continually increasing; much of such guarding is entitled to a credit under merit rating. The average condition of plants from point of safety is therefore continually improving. A schedule in balance at the starting point would soon be out of balance because the foundation upon which it was built, the average condition, has changed. Credits in excess of charges would necessarily be the result. This could, of course, be overcome by erecting a schedule upon the basis of a perfect condition and giving charges only. This would not result in a balanced plan, but would take care of changing general or average conditions. The present circumstances making such a schedule impracticable have been set forth in Dr. Downey's paper.

The point was also made that improved safety conditions over the average assumed are gradually reflected in the pure premiums, and that a double credit results by giving reductions in rate for certain items. This is no doubt an important fact, and is a weighty argument for periodic revisions of the schedule, raising the standards and eliminating credits.

As pointed out by Dr. Downey, one of the most valuable features accomplished, at least in part, by the Industrial Compensation Rating Schedule is the adjustment of the standards to conform

more nearly to the legal standards of the various states. Many controversies have arisen in Wisconsin because of the difference in the standards of the state and the insurance carriers. The Wisconsin Industrial Commission has vigorously inspected elevators for many years. Few elevators in the state do not meet the commission's requirements. The standards in the 1916 schedule were, in some instances, more stringent than the state standard, and the employer who had fully met the state standard, acting under instructions from state inspectors, and then was charged for not meeting the rating standard, objected to paying the charge, and did so with justice. To overcome the difficulty the rating standards on elevators were changed to conform to the state standards. State authorities are anxious for unified standards and are open to conviction regarding the correctness of their requirements. This good work should be continued; it is the greatest advance for proper guarding, and when enough states will have adopted the common standard manufacturers of machinery will be compelled by competition to make their products complete, including the universal guard.

In discussing the 1918 schedule Dr. Downey compares the method of basing credits and charges upon the number of employes with the old plan of flat charges and credits. The inaccurate method of reducing items to cents on the rate, making necessary the use of the estimated payroll, has been eliminated. The old evil could, of course, have been overcome by a flat addition or subtraction of the charges or credits instead of a reduction to cents on the rate.

It is extremely questionable if the number of employees can be ascertained in advance with much more accuracy than the payroll. Dr. Downey suggests that the average may be obtained from plant records taken at monthly or quarterly intervals. These must necessarily be the records of the past period. So also can the payroll be determined with even greater accuracy for the past year or part of the year. The records of the past period may, however, be of little value for estimates of the future. During this period payrolls have fluctuated greatly from year to year. Such variations are likely to continue until industrial conditions have again become stabilized. The number of employees has fluctuated in about the same degree as the payroll. Under present conditions at least it does not appear that the number of employees can be more accurately estimated than the payroll. Notwithstanding this fact, if it is correct, the plan of the 1918 schedule is superior to the schedules of the past.

The success of the 1918 schedule must depend to a large degree upon the accuracy of the value of $K$ in the various formulæ. It should not be long before sufficient statistics have developed to aid in producing values with considerable accuracy. The methods employed in determining the value for $K$ were not fully explained in Dr. Downey's paper.

The unfortunate feature of schedule rating of physical hazards is that the items chiefly considered produce only about one fifth of all the accidents. During the three years 1915, 1916 and 1917, 40,980 injuries were compensated in Wisconsin. Of these only nineteen per cent. were caused by machinery, with a corresponding time loss, weighted in accordance with the recommendations of the Statistical Committee of the International Association of Industrial Accident Boards and Commissions, of less than twenty-four per cent. of the total.

MR. WILLIAM NEWELL:

The subject has been so well covered in this paper and the details so thoroughly discussed by the author, that I fear there is little of value that I can add. Having charge of the inspection work of a compensation insurance carrier naturally leads me to consider the schedule from the point of view of its practical application, with special reference to the determination of the charge removed or credit granted, expressed in dollars and cents per year, for the safeguarding of individual conditions. In common with other carriers, we receive numerous requests from our assured to whom we have sent safety recommendations, asking us to advise them the credit they will receive for compliance with each individual item, and they invariably desire to have this expressed in dollars and cents annual saving per item, together with the total saving expressed both in terms of premium reduction and rate reduction.

This leads me to a discussion of the so-called flat premium values in the schedule. Mr. Downey states on page 331, in referring to the 1916 Schedule, that "the flat values, lastly, are not in practice—what the theory requires—fixed premium amounts," pointing out that the rate as respects the flat values necessarily depends upon the payroll disclosed by the policy declaration, differences in payroll estimate causing these values to fluctuate by as much as one hundred per cent. This was overcome to some extent by the Compensation Inspection Rating Board of New York by using an estimated payroll based on the average number of employees during the policy year as determined by inspection and the approximate annual average wage in the industry classification. In the footnote on page 331 Mr. Downey states "In the 1916 Pennsylvania Schedule an attempt was made to treat the flat charges as net premium additions, over the above the rate expressed in the policy. This practice would have realized the theory of the flat items—only the flat premium additions proved uncollectible in practice." I quote this here because I shall have occasion to refer back to it presently.

Mr. Downey states in Section 2 on page 332 that the Committee wholly abandoned the flat premium method of value expression in the 1918 Schedule, all items being valued either in per cent. of manual rate or in cents on payroll. This is true as regards the values carried into the charge and credit columns on either side of

the rating form, but I am unable to see that such is the case in the assignment of flat charge and credit points under Schedule items 120 to 210 inclusive, comprising defective flooring, floor openings, stairs, elevated runways, traveling cranes, elevators, and certain boiler and electrical defects, which are converted into terms of rate by the formula:

$$\frac{D}{E} \times K = \text{cents on pay roll},$$

in which $D$ = number of defect points multiplied by the item weight of each, $E$ = number of plant employees, $K$ = a constant per unit of payroll.

The author states on page 337 that this formula gives a practically constant premium value per defect, and that pending the statistical determination of $D$, the actual judgment values of the former schedule were retained. If $D$ is to represent these values in dollars as well as in points the formula becomes:

$$\frac{D}{E} \times 12.5 = \text{cents on pay roll},$$

which is equivalent to assuming an average annual wage of $800 per employee. The value of a point will only be exactly a dollar, however, in manual classifications in which the payroll averages $800 per employee per year, and for any other classifications the value of a point will bear the same ratio to a dollar as the average annual wage bears to $800. Thus an assured whose payroll per employee averages $800 per year will be charged $2.00 annually for a non-standard elevator gate, while a risk averaging $600 will be charged only 75 per cent. of $2.00, or $1.50, and a risk averaging $1,000 will be charged 125 per cent. of $2.00, or $2.50. This comes about through the fact that although the charge for items 120 to 210 inclusive (or credit in case of superior conditions) expressed in cents on payroll by use of the above formula will be constant for all risks having the same total of defect points and same number of employees, the premium derived by applying this rate increase to the actual total payroll will naturally vary directly with the average payroll per employee.

The author states on page 335: " Given the same number of employees, an unrailed stair, an unguarded elevator entrance or a hole in the floor presents practically the same risk of injury in a silk factory as in a carpenter shop." The actual premium charges for the defects mentioned would not be the same, however, for the reason explained above—the premium charge per defect in the silk mill would bear the same ratio to the premium charge per similar defect in the carpenter shop as the average annual payroll per employee in the former bears to the payroll per employee in the latter.

This does not coincide with the theory expressed by Mr. Downey in his reference to attempting to collect the flat charges as net premium additions, which I have quoted above. This latter method, which was found impracticable would of course result in making the values of the so-called flat charges and credits uniform for all risks.