Outline of a Method for Determining Basic Pure Premiums.

BY

ARNE FISHER.

"We assert that what is really statistical work must be undertaken only by the adequately trained statistician and that when it is not, then the investigation cannot be considered as falling into the field of science."

(Karl Pearson, *Biometrika,* Vol. X, page 172.)

I. Introductory Remarks.

In reading over the first volume of the *Proceedings* of our Society, one will readily notice that the subject of social insurance in its various forms and aspects is still in its infancy in this country. Our newly formed organization has, however, already done valuable service in bringing together and crystalizing a number of problems under this particular branch of insurance. In fact, the majority of papers read at previous meetings have been in the nature of stating rather than of solving problems. The technique of our science is as yet in a somewhat crude state, and the splendid attempts of several members of this body to build up a theory on a sort of mathematical basis, I think, are as yet in an essentially empirical state and do not constitute the final word in the matter. I do not make this statement in a spirit of hostile criticism as I readily realize that dire necessity rather than lack of mathematical knowledge have driven such investigators to adopt essentially empirical formulas and methods. The passing of workmens' compensation acts in a number of states during the last few years has put a vast project before the American insurance world and makes in the complete absence of compact data upon which to construct reliable and just rates, the efforts of those gentlemen highly commendable. They have, in fact, done pioneer work in clearing the forest of the trees and given us a clear view of the lay of the land. A lot of snaggy brush and undergrowth remains still in the ground, however, and must be removed before we can start to till the virgin soil in a systematic manner. In clearing the soil of such snags we must look out for the future and make the task for the plowman as easy as possible by leaving no holes and pitfalls in the cleared ground. The time has about arrived at which the proposed methods ought to be advanced enough to be subjected to the test of a critical analysis and systematic statistical methods substituted for the preliminary empirical rules and made the basis for future research work in social insurance.

The investigation of causal relation between various social and economic phenomena presents many pitfalls to the uninitiated and offers many opportunities for fallacious conclusions. The statistician in investigating the effect of a multitude of causes working on a single object or group of objects as to a certain attribute cannot narrow himself down to the relatively simple issues of the physicist and the chemist, who in most cases are able to investigate the variations as to a certain attribute from a single cause. In view of the complexity of the problem of social insurance it will, I think, generally be allowed that the statistical methods hitherto employed are frequently inadequate. For this reason, I deem no apology necessary in presenting a method that, as far as uniformity and general systematic procedure is concerned, exceeds any other covering the same ground.

It is my opinion that in the solution of the problems of social insurance we ought to follow the modern statistical methods of the English biometricians and continental statisticians, especially the Scandinavians, rather than the old methods put forward by life insurance actuaries. Your future work falls in quite a different field of statistical science than the work of an actuary of a life insurance company. The task of the latter falls essentially inside that division of statistical methods, which Charlier has called homograde statistics (alternative statistics). In life insurance we deal mainly with two alternative and easily distinguishable contingencies, life or death, which allow a sharp and easy division by a dichotomous judgment. As far as the actuary is concerned there is no such thing as a half dead person or even a half shot person. When turning to the subject of invalidity, we encounter, however, quite a different state of affairs. We may of course easily distinguish between active healthy persons and disabled persons, but the latter do by no means present a compact homogenous group. The degree or intensity of invalidity is a variable factor and belongs to the second great division of statistical methods, the heterograde statistical series. It is in this latter branch that during the last ten to fifteen years we have witnessed the splendid development under the able and energetic leadership under such men as Pearson, Elderton, Edgeworth and Yule of England, Westergaard, Charlier and Kjaer of Scandinavia, Tschuproff and Nekrasoff of Russia and Kapteyn of Holland and a host of other brilliant investigators.

II. The Test of the Stability of Statistical Series.

Statistical science or methods as I view them fall in two parts: the collection of data and the analysis of such data by means of mathematical-statistical methods. It is with the latter part I will deal exclusively in this paper. A mere collection of statistical data, as for instance the data in a census report, means little beyond giving a faithful picture of the past and present conditions of a certain community. We would, however, be very wrong in assuming with Stuart Mill that such picture would remain essentially the same in the future and draw conclusions as to the future by mere induction. It is quite true that many statistical series remain practically unaltered both as to time and to space, but the majority are subject to fluctuations. Before using collected statistical series as the basis of predictions of future events, we must first of all investigate the stability of such series.

Here in America the starting point in compensation premiums or rate making has been a classification pure premium defined by the symbolic formula: $\pi = L/P$ (Losses $\div$ Payroll). When observations of this kind are published it is seldom necessary or even convenient to publish the primary lists in their original form, but the whole experience (population) is divided into sets (samples) of payrolls consisting of say

$$P_1, P_2, P_3, \cdots P_N$$

monetary items. Now let in the same manner

$$L_1, L_2, L_3, \cdots L_N$$

represent the corresponding losses to each of the observed sets. We then have a certain kind of statistical series:

$$L_1, L_2, L_3, \cdots L_N$$

which is called a *homograde* statistical series. The members L_1, L_2, L_3, $\cdots$ L_N are called the *elements* of the series and the correlated numbers P_1, P_2, P_3, $\cdots$ P_N are called the *numbers of comparison.*

It lies now close at hand to regard the ratio L/P inside a singular classification in the light of a constant mathematical probability whose numerator expresses the unfavorable events (losses) while the denominator gives the total possible cases (total payroll). If such actually was the case, then the series $L_1, L_2, L_3, \cdots L_N$ of the various sample sets inside a particular classification would represent what in homograde statistics is called a *Bernoullian Series,* corresponding to a series of samples of drawings of variously colored balls from an urn. In such series we could a priori by means of a simple mathematical formula predict the magnitude of the fluctuations from the average value from set to set. The older school of statisticians regarded such an identity between a statistical series and a Bernoullian Series almost as an axiom, and even many modern statisticians hold a similar view. As far as I can judge the previous writers on the subject both in the *Transactions* of the Actuarial Society of America and also in the *Proceedings* seem to have taken the same attitude, although they may not be ready to admit so off hand. Is this attitude correct? Or in other words, have we any reasons for assuming an absolute identity between a Bernoullian Series and an actual series in compensation insurance? A priori there exists no earthly reason for such a view except a somewhat mythical belief in the vague term of "the law of large numbers." Somehow or other the belief has crept in that if the exposure is large enough, then the pure premium will also be sufficient. This of course would be true if we were actually dealing with a Bernoullian Series, but as we know all statistical series—even at the best—are approximations only to the Bernoullian form of series. The only way to determine the existence of even an approximate identity between the loss series $L_1, L_2, L_3, \cdots L_N$ and a Bernoullian Series and to establish the stability of pure premiums is by an actual test, and as far as I know no such test has even been attempted in America.

The test of stability is easily made by means of the Lexian-Charlier Theory of Dispersion. Taking the observed losses of various establishments inside the same classification we may find a measure of the magnitude of actual fluctuations inside such classifications by calculating the coefficient of disturbancy. The method is simple and speedy and ought to be applied in definitely establishing the presence or non-presence of stability from year to year or from plant to plant. I am by no means too sure that the test will establish a state of stability even inside kindred industrial hazards. Thus we know with absolute certainty that the hazard as to fatal accidents in the mining industry by no means is stable and uniform for the whole United States but varies greatly from state to state. On the other hand, states like Iowa and Michigan exhibit actual mathematical imaginary coefficients of disturbancy, indicating very stable conditions from year to year.

III. The Application of the Theory of Dispersion in Dividing the Material in Proper Classifications.

So far I have only discussed the application of the theory of dispersion in testing the stability of the loss series inside a particular classification. I shall next briefly outline how the same method may be used in grouping classifications. As Mr. Mowbray and Dr. Downey have pointed out in their recent contributions, the classifications as to products as employed in most of our previous experiences and state industrial commissions means little or nothing in obtaining a proper estimate of the hazard of an insured object. Take for instance such vast industries as those of the automobile, textile and shoe industries. Each such industrial plant represents a little world all by itself, and it is readily seen that in so far as individual hazards are concerned—or even group hazards of several individuals—the employees represent a decidedly heterogenous mass. From a cursory view it would appear that a workman in the automobile factory foundry is exposed to much greater risks than his fellow worker in the upholstering department. Yet, on the other hand, it seems very plausible that workmen in different industrial plants, performing quite different duties and sorts of labor, nevertheless from the underwriters' point of view present the same insurance risk, and it is thus possible to group workmen from certain parts of the shoe factory with workmen from certain parts of the automobile factory.

It is now the task of the statistician to pick out the individual workmen inside the various plants and group them together in similar classifications of hazards. Once having agreed on the structure of such a group of individuals making a certain classification, we may test its stability by means of the theory of dispersion. If the coefficient of disturbancy is either zero or an imaginary quantity, we know the series is stable. If, on the other hand, it has a large positive value, then we know that the classification is not homogeneous and will in most cases resort to a further sub-division.

In making such a classification of hazards, I think, it is essential that we limit ourselves to a comparatively small number of classifications, say possibly fifty or sixty. I am quite well aware of the fact that the famous Massachusetts schedule *Z* contains over 1,000 classifications. With due respect for the eminent statisticians and actuaries who have made up this schedule, I dare nevertheless express as my personal opinion that the authors of schedule *Z* have made a sort of classification which is the very antithesis of scientific statistical analysis. It seems to me that these gentlemen have attempted to investigate the variations due from a single cause instead of variations due from a multitude of various causes, the paramount object of statistical methods. In making such a vast classification they have on the other hand gotten stones instead of bread, as it is evident that a very large number of the various classifications will contain such a small number of exposures or numbers of comparison (in this case the amount of the payrolls) that the pure premiums are affected with so enormous errors of sampling that they are absolutely unreliable.

A limited number of classifications is desirable from purely practical reasons and later on for the graduation of pure premiums. I shall take the liberty to illustrate this statement by transgressing on the question of sub-standard risks in life insurance as developed by certain Scandinavian and Austrian actuaries. These actuaries emphasize the importance of a limited number of danger classes and divide the sub-standard applicants into three danger classes for which they construct separate mortality tables with special premiums and reserves. The examining physician simply makes a detailed diagnosis along the lines indicated in the questions on the application blank. The diagnosis enables the actuary to place the individual applicant in the proper danger class. When it is possible to throw the various shades of sub-standard risks into three great classes, I deem it quite possible to limit the number of classifications of industrial hazards to the size as suggested above.

In making such a classification by purely statistical methods I quite readily realize the enormous obstacles which even the most highly trained statistician would find difficult to overcome in the preliminary work in making the first, so to say, rough picking of the various elements. No single statistical expert, no matter how well trained, has such a detailed knowledge of the numerous and varied industries and trades in America that he is enabled a priori to make such a classification. The number of possible ways into which the various degrees of hazards might be permuted into classifications would theoretically approach infinity and would therefore actually exclude trial classifications taken at random without subjective information as to the character of the risk or hazard in question. I therefore quite agree with Mr. Mowbray's* suggestion to work in harmony with the safety engineer in making a preliminary classification. But I can not follow the same author when he tries to extend his method of personal judgment in rate making in a rather arbitrary manner. Poisson in his celebrated treatise on "Recherches sur les probabilites des jugesments" has shown how carefully we must act when relying solely upon personal judgment and subjective information. Moreover the method of "weighting," if it can be called so, by Mr. Mowbray and the proposed modification by Dr. Rubinow shows a so complete disregard of errors due to sampling that it by no means can be considered an improvement over the original judgment percentages of the unit classification rating.

Such arbitrary methods are eliminated when we make use of the dispersion theory in a plan which I shall now proceed to outline. The safety engineers, the inspectors of industrial plants, the owners, superintendents, foremen, etc., of such plants in conjunction with the mathematically trained statistician make first a preliminary survey of the various trades and industries and group the workmen thereunder in some of these fifty above mentioned danger classes. The payrolls are divided accordingly. Also the losses, after being reduced to the same standard of comparison, a process which I shall describe shortly below.

Let

$$\begin{gathered} L_1,\ L_2,\ L_3,\ \cdots\ L_N \\ P_1,\ P_2,\ P_3,\ \cdots\ P_N \end{gathered}$$

be such samples of payrolls and losses from various trades and industries which according to the mutual judgment of the safety engineer and the statistician represent the same degree of hazard. In such a homograde series the sample P_1 may possibly stand for a section of the payroll from the glass industry, P_2 a section from the

* *Proceedings*, Vol. II, p. 129.

textile industry, P_3 from the lumber industry, and so forth for the total N samples which according to personal judgment should belong to the same classification and possess the same degree of hazard. Now it would indeed be rash to assume that a first selection would hit the true mark and obtain a classification of perfectly homogenous hazards where possible variations from industry to industry were entirely due to errors of sampling (variations from chance). Hence it would be fallacious to club all the samples together and form a pure premium as

$$\Sigma L_i/\Sigma P_i \quad (i=1, 2, 3, \cdots N),$$

as we in this manner would rely solely upon a subjective judgment. By keeping the original N sample sets separate and applying the mathematical criteria from the theory of dispersion, we may test the stability of the classification and in this way get a purely quantitative measure of the soundness of the engineer's judgment. In this connection I may mention that I have developed a few mathematical criteria entailing the use of the higher statistical characteristics: the coefficient of disturbancy, the coefficient of dependency and the coefficient of variance by which I am able to test not alone the stability of the classification but also to tell in what direction —positive or negative—the safety expert has made errors in estimating the various hazards. As is well known from experimental psychology, subjective judgments of this kind are generally affected with systematic errors.

This plan is similar to that of Mr. Mowbray in making a partial use of subjective judgment. But whereas Mr. Mowbray apparently makes such judgments the basis for actual rate making, the method as suggested above only uses the subjective judgment of the engineering expert in making the first rough picking of the hazardous elements in the homograde series. The test of stability is then done by purely deductive methods and gives us a criterion as to the validity of the engineering judgment in a purely objective manner. Objective statistical methods are thus in the final analysis made the paramount issue and play a much greater role than in Mr. Mowbray's method.

In this manner we obtain an absolutely safe statistical test as to the uniformity and homogeneity of judgment classification of the safety engineer. But right here I wish to emphasize that it would be advisable to let the dry and dispassionate statistician act as a brake on the sometimes very fertile imagination and enthusiasm of our "safety first" engineers and propagandists. When the safety engineer actually begins to differentiate between an elevator accident and an accident due to slippery floors in entering the elevator, he has indeed made a most serious slip himself in so far as the statistical analysis is concerned, and he has completely slipped away from purely objective research and soared to heights and regions far beyond the horizon of the prosaic statist. If such hair splitting distinctions—mere apples of Eris for the engineer and apples of Sodom for the statistician—must be made the basis for compensation rates, we statisticians may as well pack our kits and leave the field to the undisputed possession of the engineer and "efficiency" expert. In this case I see practically no limits to which such sophistries might be carried. For instance, it would be but a short step to differentiate between floors slippery from oil, butter, axle grease, lemon juice, etc. Even I, as a most dry statistician, have enough poetic imagination to appreciate the new and vast vistas laid open to speculative thoughts by a fertile brain by this method, and I would indeed not be surprised some day to read a serious and learned discussion between safety engineers on the various degrees of hazards of slipping on a banana or a lemon peel. Undoubtedly the safety expert takes himself seriously and regards himself as an essentially practical man. To my simple notion (and here I speak as a former engineer) such finesses appear pedantic rather than practical, and—paradoxically as it may appear—certain safety devices ought to be applied to the safety expert if we really are to follow the "safety first" motto and not drift into idle and extravagant speculations. As a former engineer I quite well appreciate the temptations for the engineer, trained in essentially experimental sciences as physics and chemistry, to try to trace the effect of a single cause. In my early youth, before I obtained a clear understanding, of statistical methods, I held similar views, and first after actual statistical experience I realized that the domain of statistics was the investigation of the effects of a multitude of causes instead of the effect of a single cause on a particular object.

As to the record of losses it must be borne in mind that those items must be reduced to the same standard of comparison. This will of course necessitate a reduction in many cases where the compensation benefits vary according to the different states. I have been assured by several casualty insurance officials with whom I have discussed this matter that such reduction presents no serious difficulties.

A word or two about mean or average rates is perhaps not out of place in this discussion. Somehow or other the belief has crept in, in many statistical investigations, that when the exposure is large we will always obtain a safe "average rate." It must, however, be borne in mind that a large exposure is gained only by sacrificing the homogeneity of the material. The mean value is but a particular value of a large number of variates. If the range of variation is large and fluctuations are marked inside that range, the mean value does by no means present the most probable or even the expected value. It is therefore of the utmost importance to study the fluctuations from the mean from sample to sample. This is exactly what is done in the theory of dispersion. By breaking the large mass up into several samples and computing the higher statistical characteristics of such samples we obtain a safe measure of the dispersion or spreading of the various samples about the mean. In this manner it frequently happens we get much more reliable results out of a small homogenous total exposure than a large heterogenous exposure.

In submitting the above plan for the construction of pure premiums, I do not wish to make any conjectures as to the probable outcome of the test of stability. Possibly the numerical results of the calculation might be of negative value in so far as they might establish that the series was unstable. Such a result—negative as it might be—would, however, be of value in calling our attention to needed reforms of the very principles of rate making. Without making any predictions whatsoever, I think, nevertheless, that the dispersion will establish essentially stable conditions inside homogenous classifications.

The importance and value of the outlined scheme lies, however, in its general application and systematic procedure and in choosing for its basis a purely objective method and analysis in the place of vague subjective judgments and arbitrary rules. Moreover, the method as proposed is built upon an absolutely sound and valid mathematical foundation as developed by the world's foremost experts on statistical methods, which is more than can be said about the methods hitherto suggested in rate making in social insurance in this country.

NOTE.

Since writing this paper I have discussed the matter contained therein with a few casualty statisticians. These gentlemen—although recognizing the purely theoretical value of the method—seem to be of the opinion that it is not feasible to split up the payrolls in the various industries according to the method as outlined above, and they have suggested that I outline a procedure as to the collection of the data necessary for a mathematical analysis. Although my paper dealt exclusively with the analysis of the collected data, I shall nevertheless take the liberty to submit the following short remarks as to the construction of the primary lists from which the homograde series must be constructed.

To begin with, let us consider a certain large and typical industry, such as the automobile industry. The safety engineer, in conjunction with other experts, assigns each individual hazard as represented by the individual employee to one of a limited number of classifications. Each employee is thus assigned to a certain danger class according to the purely subjective judgment of the safety experts and the payrolls are divided accordingly. The losses accruing during the year are likewise distributed among their proper classifications of danger classes.

In this manner we will have N classifications inside the automobile industry (Industry No. 1) with payrolls and losses arranged in the following homograde series:

$$\left.\begin{matrix} {}^{(1)}L_1, & {}^{(1)}L_2, & {}^{(1)}L_3, & \cdots & {}^{(1)}L_N, \\ {}^{(1)}P_1, & {}^{(1)}P_2, & {}^{(1)}P_3, & \cdots & {}^{(1)}P_N, \end{matrix}\right\} \text{No. I.}$$

Again let us take another industry, say the shoe industry (Industry No. 2). Here we also let the safety experts make their classification according to hazards and obtain M classifications of losses and payrolls arranged in the following homograde series:

$$\left.\begin{matrix} {}^{(2)}L_1, & {}^{(2)}L_2, & {}^{(2)}L_3, & \cdots & {}^{(2)}L_M, \\ {}^{(2)}P_1, & {}^{(2)}P_2, & {}^{(2)}P_3, & \cdots & {}^{(2)}P_M, \end{matrix}\right\} \text{No. II.}$$

A third industry, for example the building trades, may yield the following series of R classifications:

$$\left.\begin{matrix} {}^{(3)}L_1, & {}^{(3)}L_2, & {}^{(3)}L_3, & \cdots & {}^{(3)}L_R, \\ {}^{(3)}P_1, & {}^{(3)}P_2, & {}^{(3)}P_3, & \cdots & {}^{(3)}P_R, \end{matrix}\right\} \text{No. III.}$$

Continuing in the same manner we may easily obtain fifty or more typical sample classifications from various industries.

These primary lists are now at the completion of the calendar year submitted to the statistician. The losses are reduced to a common hypothetical standard of payroll P. Such reduction requires in most cases a certain correction on account of errors from sampling, which, however, may be taken care of by applying certain well-known criteria from the theory of probabilities.

After having reduced the loss series to a common standard of payroll we obtain the following *reduced* homograde series:*

$$\begin{array}{lllll}
{}^{(1)}L_1, & {}^{(1)}L_2, & {}^{(1)}L_3, & {}^{(1)}L_4, & \cdots, \\
{}^{(2)}L_1, & {}^{(2)}L_2, & {}^{(2)}L_3, & {}^{(2)}L_4, & \cdots, \\
{}^{(3)}L_1, & {}^{(3)}L_2, & {}^{(3)}L_3, & {}^{(3)}L_4, & \cdots, \\
\cdots & \cdots & \cdots & \cdots & \cdots, \\
{}^{(50)}L_1, & {}^{(50)}L_2, & {}^{(50)}L_3, & {}^{(50)}L_4, & \cdots.
\end{array}$$

It is now the statistician's task to pick out and combine such values of the various L's so that we shall finally have a limited number, say possibly sixty or less classifications of hazards.

Let one of those picked classifications contain the following samples of L:

$${}^{(1)}L_{10}, \ {}^{(2)}L_5, \ {}^{(3)}L_{24}, \ {}^{(4)}L_9, \ \cdots \ {}^{(50)}L_{12}.$$

This means that ${}^{(1)}L_{10}$ (classification No. 10 of the automobile industry, Industry No. 1), classification ${}^{(2)}L_5$ (classification No. 5 of the shoe industry, Industry No. 2) and classification ${}^{(3)}L_{24}$ of the building trades and so forth for the various industries, should represent the same hazard. This series is now tested as to stability. If the mathematical criteria are satisfied, the series may be considered as stable and the pure premium expressed as

$$\pi = \frac{\Sigma L}{\Sigma P}.$$

If the criteria are not satisfied, it becomes necessary to resort to another grouping of the above reduced series of L. However, any adequately trained statistician will not need to make very many trials before he will be able to produce a stable series, that is, a series which may be considered stable for practical purposes.

* Such a homograde reduced series is as a matter of fact a series of pure premiums for the various sub-classifications.

In conclusion, I wish to state that if any of the members of this Society should feel inclined to undertake a research along the lines indicated in this paper the only requisite is a thorough understanding of the mathematical theory underlying modern statistical methods. *This requisite is essential, however, and without it no valid results will be reached. I might also add that although I have treated the rudiments of the dispersion theory in my book on "The Mathematical Theory of Probabilities" the treatment there is not adequate in this case where certain higher statistical parameters (such as the eccentricity and the excess) must be taken into consideration.* The actual arithmetical computations are quite simple and may be undertaken by any office clerk.