JIA 75 (1949) 0183-0231

SOME THOUGHTS ON THE ANALYSIS OF NUMERICAL DATA

BY L. G. K. STARKE, B.A., F.I.A.
of the Government Actuary's Department

[Submitted to the Institute, 28 March 1949]

THIS paper is the record of an attempt by an ordinary actuary, who entered the profession in days when statistics were but a nugatory ingredient in the examination syllabus, to get a little clearer in his mind about some of the similarities and differences between the traditional technique of the actuary and the methods which have been developed for dealing with statistical material in other fields. It seemed that the result might be of some general interest, and it is from this point of view that the paper is submitted to the Institute.

2. The interests and responsibilities of the actuary have expanded a great deal since the early days of the profession. It is likely that this expansion will continue; but nothing that the future may hold in store can alter the fact that the origins of actuarial thought and practice lie in the study of life contingencies. This study has never been simply an intellectual exercise performed in the hope of increasing the sum of human knowledge, or bettering the lot of mankind in some general kind of way; it has been pursued for the specific purpose of providing a systematic basis for the sale of life assurances and annuities. There are many other fields of inquiry in which methods of dealing with numerical data have been developed with some clearly defined purpose in view. At the same time, since figures are an inevitable by-product of the conduct of affairs in a civilized community, it is natural—man being an inquisitive creature—that much work should also have been done in devising systematic methods of sifting and analysing collections of statistical material of all kinds, rather in the hope of obtaining ideas, or of giving a more definite shape to ideas already formulated in general terms, than with a limited and precise objective. That actuaries have played their part in exploratory activities of this sort does not alter the historical fact stated above.*

3. To regard the whole field of numerical-mathematical activity as sharply divisible between strictly purposive calculations and tentative researches into the unknown would oversimplify the picture. Actually, a good deal of statistical research is undertaken in the knowledge that, if it produces positive results, there are practical uses to which those results can at once be applied. But the broad distinction serves a useful purpose as an introduction to the general proposition that the method of approach to a problem which involves the use of statistics is governed to a very large extent by two factors: the nature of the operator's objective and the extent to which he is free to select and adapt his raw material. Other considerations may come in, but these two provide a sufficient basis on which to compare the behaviour of an actuary embarking on the classical exercise of constructing a mortality table from assurance records with that of an investigator in the social or economic field who seeks to interpret one statistical time-series in terms of others. This comparison is not invalidated by the circumstance that the investigation of the mortality of assured lives is nowadays a centralized and co-operative enterprise. The fact remains that actuarial science grew out of the experiments of individuals in this field, and it continues to form the hard core of our educational curriculum.

* See note on p. 217.

ACTUARIAL OBJECTIVE AND METHOD

4. Let us consider first the position of the actuary in relation to his data. In the organization which collects these he is a responsible official whose ideas and requirements are treated with respect—probably, indeed, with something approaching reverence. He may actually be in charge of the whole organization. Further, his company's desire, for business reasons, that the information supplied by prospective clients shall be as complete and accurate as possible is fully in accord with his own predilections as a technician; and the prospective clients supply the information in the knowledge that if it is subsequently discovered to be materially incorrect the contract which they are seeking to make may be declared null and void. Human nature is fallible, and it would be going too far to suggest that the raw material of an investigation into the mortality of assured lives is as free from error as measurements of animal, vegetable or mineral matter performed with delicate instruments in a laboratory; but its reliability ranking in any catalogue of statistics derived from the behaviour of individual men and women must be very high.

5. Equally propitious is the fact that any one of the units of which the statistical aggregate is comprised can be identified and detached at will, so that the number of ways in which the data can be sorted and classified is limited only by the number of separate characteristics noted in the individual record. Expressed mathematically, this means that the actuary can go a very long way in the direction of isolating his variables; this circumstance gives him an enormous technical advantage over many statisticians and is, I think, one of the fundamental reasons why the methods of actuarial inquiry have developed along lines which differ so considerably from those which have been evolved to suit the statistical facilities available in other fields of investigation.

6. The idea that, generally speaking, old people are more likely to die than young people must have been common property long before mortality tables were thought of; to trace its origin we should have to go back to the stage in pre-history at which primitive man began to use his powers of observation as a basis for reasoning. Just when it was first realized that the female of the species is hardier than the male I do not know; the discovery probably came very much later. At any rate, it is safe to say that for the primary purpose of constructing monetary bases for the sale of life assurances and annuities, actuaries everywhere have long since founded their calculations on the axiom that the principal personal characteristics in relation to which mortality should be studied are age and sex. The standard method of exhibiting the results of a mortality investigation is by a one-dimensional table of death-rates for each sex, with age as the argument.

7. This does not mean that the actuary is blind to the existence of other variables; but if the contrast suggested in paragraph 3 is to be fully established, attention must be drawn at this point to certain aspects of his attitude to these other variables. In performing his analysis he has regard mainly, if not entirely, to those potential influences upon mortality which can be sought for within the experience which he is investigating—that is to say, those which can be identified by reference to the recorded particulars of each individual member of the body of assured lives under examination. When the particular mortality table which he is constructing is in its final shape, it can of course be considered in relation to the numerous economic, social or other circumstances which may have a bearing upon the level of death-rates in general; but this is

a process separate from, and subsequent to, the technical operation of constructing the table itself. It is, indeed, not easy to see how the two processes could be combined if the cardinal principle of treating age and sex as the predominant characteristics is to be retained. In considering the mortality table against the more general background just mentioned, the usual practice is to compare it with other tables, constructed in accordance with the same conventions, which relate to other classes of lives or other periods of time; but this comparison is qualitative rather than quantitative in the sense that it does not usually involve any considerable attempt at formal analysis of the statistical end-product as distinct from the statistical raw material.

8. The treatment of the additional variables which are discoverable within the experience itself is limited, in the first place, by the fact that few of the recorded characteristics of assured lives other than age and duration of assurance can be expressed numerically. Secondly, even if they could be so expressed, the volume of data may not be large enough to permit of analysis by age and some other variable simultaneously. Thirdly, if neither of these limitations existed, the exhibition of rates of mortality on a multivariate basis would be seriously hampered by our physical inability to construct tables in more than two dimensions.

9. How, then, does the actuary proceed? Here it is relevant to quote from Hardy's First Lecture:

> In the selection of data suitable for his purpose the Actuary will aim on the one hand at a sufficiently broad basis both in space and time to eliminate the effects of local and temporary fluctuations, and on the other hand he will aim at obtaining as far as possible a homogeneous group of data. These two aims are more or less in conflict, and he will lean to the one side or the other, according to the object he has in view. Where, for example, that object is to produce a table that may be adopted as a general standard by various institutions, often differing considerably as to their individual experience, he must aim at a correspondingly broad foundation. In these circumstances it will not generally be possible to obtain a really homogeneous experience. If it is a question of the mortality of assured lives, for instance, this will be found to be affected by endless individual variations, age, sex, duration of assurance, occupation, civil condition, class of assurance, character of the insuring office, etc., etc., and from such material approximately homogeneous data could only be obtained by cutting up the experience into comparatively small groups and thus sacrificing all generality. This can be avoided in practice by first excluding all extreme variations. The sexes will be separately treated, lives so impaired as to prospects of longevity by personal health, family history, occupation, or residence in unhealthy districts as to be 'rated up' will be excluded, as also classes of assurance that may be supposed subject to rates of mortality differing from the average. When the data has thus been trimmed of the extreme variations, a body of experience will generally remain not greatly shrunken from its original dimensions and in which the discontinuous variations are sufficiently numerous and individually unimportant to render the data for practical purposes homogeneous. The rates of mortality, or of withdrawal, can then be treated as functions of the two remaining variables of importance, the age and the time elapsed from date of entry; or as functions of the age only from the point at which the factor of duration may be found to be unimportant.

10. It seems fair, therefore, to state that in his traditional capacity as a constructor of mortality tables the actuary is not beset, and never has been beset, by the problems of multivariate analysis which play so large a part in other forms of statistical investigation, for the reasons that (*a*) the recognition of age as the principal factor determining the mortality rate is as old as actuarial

science itself, and (*b*) the actuary enjoys the enormous advantage of being able to dispose of the effect of secondary variables by sorting and pruning his material before the individual units of which it is composed have lost their identity by being merged into statistical aggregates. His aggregates are, in fact, just what he likes to make them in the course of his search for the 'homogeneity' which Hardy has described in the passage quoted above. The experimental probing and dissection of such aggregates by mathematical methods in an attempt to identify the factors which have combined to produce them and to assess the relative importance of these factors—in short, the process of analysis as commonly understood in other fields of statistical work—hardly come into the picture.

11. The foregoing remarks might appear to suggest to anyone outside the actuarial profession that the construction of a mortality table is a simple exercise for the intelligent schoolboy rather than a skilled operation which has merited the attention of trained and matured minds for something like two hundred years. Let me therefore say at once that I have deliberately simplified my description for the sake of emphasizing the contrast which I am seeking to establish. In a paper submitted for discussion by other members of the Institute it ought not to be necessary to dilate upon the intricacies of 'exposure to risk'—a conception and a technique which emerged from the study of life contingencies, or to point out that the operation of separating off the heterogeneous elements in an experience of assured lives is not just a matter of taste and fancy but a process which may involve a nice appreciation of the value of significance tests. It may, however, be suggested in passing that the principle of exposure to risk could never have been developed if the necessary details in respect of every individual unit entering into the experience were not available. To underline this point it is only necessary to mention the apparently insoluble problem of deriving a suitable 'exposed to risk' for the realistic analysis of road accident statistics.

12. For the same reason, no reference has been made to the subject of graduation. Here again, in the development of summation formulae, for example, the actuary has made a unique and ingenious contribution to numerical mathematics; but the mere fact that so much attention has been devoted to the construction of these formulae is a sufficient reminder that the discovery of a mathematical expression enabling the rate of mortality at any age to be calculated in terms of that age is not an indispensable part* of a mortality investigation conducted for commercial purposes. Those purposes require no more than that the derived monetary functions shall progress from age to age in the general manner, and with the general smoothness and consistency, which the doctrine of life contingencies leads us to expect; and as, in any experience which has been rendered homogeneous in the manner already described, it is a necessary corollary that irregularities in the progression of the ungraduated mortality rates represent chance deviations rather than systematic features which ought to be retained in the graduated table, there can be little to choose between one smoothing process and another so long as our objective is the specific and practical one of obtaining a basis for the calculation of rates of premium, policy-values and the like.

13. The actuary would, however, be a strangely limited creature if he were

* Although it may have certain empirical advantages, e.g. in the computation of joint-life functions.

content to confine his mathematics to the derivation of purely empirical devices such as the summation formula. It is a long time now since Gompertz enunciated the doctrine of man's continuously increasing inability to avoid destruction, and since Makeham added the A to the original Bc^x and suggested further refinements. We have come to realize that the curve of the mortality rate plotted against age is a somewhat more complicated affair than these early hypotheses imply, and the tendency nowadays is to suggest that we might learn more about the nature of mortality by studying the curve of deaths than the progression of the death-rate itself; but, even as Roger Bacon and his fellows searched indefatigably for the philosopher's stone, so the hunt goes on for a mathematical relationship between mortality and age which shall hold good from 0 to ω, and no doubt a good many experiments in curve-fitting are made which never find their way into the pages of the *Journal*. The relevant point in the present context is that whether we are fitting curves to μ_x, to $\mu_x l_x$, or to E_x and θ_x separately, and whether we are searching for a 'law of mortality' or preparing a basis for a new table of premiums, we set to work instinctively along the lines which, in the early days of our science, were evolved for the investigation of mortality for financial purposes. Our main variable, age, is predetermined by tradition; we employ mathematical formulae to give precision to a shape which we already see in rough outline before us, rather than as a means of probing for we know not what; and in appraising the results of our efforts we are prone to have as much regard to their bearing on our routine professional activities as to their potential value as contributions to pure knowledge. Thus Perks was content to describe his striking developments of the Makeham curve (*J.I.A.* Vol. LXIII) as *Some Experiments in the Graduation of Mortality Statistics*, and the title of Beard's paper to the Centenary Assembly suggests that the object of his pioneer experiments in the use of the incomplete Gamma function was the approximate calculation of actuarial functions. There is perhaps something to be said for the view that while it is obviously right and proper that, as professional men, we should put the results of such experiments to the test of practical value, it is possible to be too modest about our capacity to undertake mathematical or statistical research for less obviously utilitarian motives; and that, now that mass operations in arithmetic have been so highly mechanized, devices which enable us to put the computation of derived functions on a more mathematical basis are less fittingly judged as aids to productivity than as achievements which have an intrinsic aesthetic or intellectual value.

14. It is undeniable, I think, that the purposes for which our professional technique was originally evolved, and the favourable statistical environment in which it has been practised, contrast strongly with the circumstances which prevail in other fields where calculations are made on the basis of numerical records; but the extent of the contrast, and the reasons for it, cannot easily be seen without setting up the two pictures side by side and attempting to identify the most conspicuous features of each. It is solely for this purpose that the foregoing brief crude summary of certain aspects of actuarial work has been attempted. In comparison with life-contingency work, many of the methods of general statistical analysis are infant growths; and this in itself is sufficient to suggest that what is appropriate and practicable in one case may be inappropriate or impracticable in the other. But the fact that the statistician engaged in, say, economic research often has his objective less clearly defined, and his data much less under control, than the actuary ought not to constitute a barrier

between them; nor should the fact that problems of statistical analysis outside the actuarial field are often not amenable to attack by the relatively precise methods which the actuary is accustomed to employ discourage him from bringing a disciplined mind to bear upon the possibilities of other techniques, even if they involve the disregard of time-honoured tenets. For example, in a community poor in vital statistics and devoid of actuaries a statistician who embarked upon the study of mortality as an entirely new subject would not necessarily begin by making sex and age the linchpins of his investigation.

SOME CHARACTERISTICS OF STATISTICAL ANALYSIS

15. I have chosen the comparative study of economic time-series as my contrasting picture for several reasons. The processes employed involve the use of standard statistical techniques but contain some features which seem to me to be of special interest. In two fundamental particulars—the choice of variables and the procurement of data—the position of the investigator is vastly different from that of the architect of a mortality table. The principles underlying the analysis, and the methods by which they may be applied, have recently been expounded very clearly in a monograph to which I shall refer more fully later. Lastly, in an attempt to apply these principles and methods to the interpretation of a mortality series in terms which have no reference to the characteristics of an individual life, I have encountered some difficulties which it is part of the purpose of this paper to discuss. These difficulties are described in general terms in later paragraphs; the experiment which illustrates them is presented separately in the Appendix.

16. Let us suppose that an economist wishes to discover—by statistical methods—what factors appear to be mainly responsible, and in what relative degrees, for changes from year to year in the consumption, in a particular community, of some commodity or service in general use and free supply. He will have a hunch, drawn from his knowledge of economic doctrine in general, that, of a number of factors which affect or reflect the behaviour of the community, some are almost certainly relevant, so to speak, while others are only possibly relevant, to his particular problem. How does he put this hunch to the test of analysis on the basis of quantitative data? The first part of this comprehensive question is, clearly—how does he obtain such data? Other parts of the question are, equally clearly—how does he determine the form of the mathematical expression connecting his dependent variable (the particular consumption statistic he has set out to explore) with the independent variables (the 'other factors' believed to be relevant)? How does he decide whether the inclusion, or the exclusion, of a doubtfully relevant factor will improve the fit of his relationship equation? When he has obtained what appears to be the best equation that his ingenuity and the circumstances of the case can provide, to what practical uses can it be put?

17. It appears already that questions of this kind differ considerably from those with which the compiler of a mortality experience on conventional lines need concern himself. The contrast is sharpened when we come to consider further the initial problem of getting hold of the data. There is no universal register whence our investigator can extract particulars of whatever economic or social characteristics of individual consumers—or even of a representative sample of them—he deems to be germane to his inquiry; and he may not be connected with any organization which is prepared to arrange or finance the

collection of information on this individual basis. Even if individual records were available, to convert them into statistics of the kind most likely to serve the purposes of the inquiry might be a laborious and irksome operation in view of the fact that the number and nature of the relevant factors have not been predetermined. In some kinds of statistical research access to the individual case-papers is, of course, essential; but for the sake of simplicity we will assume that in the present instance the investigator, of necessity if not from choice, procures his raw material in the shape of aggregates. Even so, the aggregates may themselves be numerous, and they may not have been published, in which event a great deal of painstaking inquiry may be needed before the quarry is run to earth in the archives of different business firms, trade organizations or Government departments. If one or more of the statistical series has to be built up piecemeal in this fashion, some differences in kind between the various components will be almost inevitable, and the whole subsequent process of seeking to establish a mathematical relationship between the series must be conducted in the knowledge that some of the quantities entering into the analysis do not represent exactly what, for the purposes of the experiment, they are supposed to represent. At the other extreme, some of the series required for the investigation may not be available in absolute terms, but only in the form of index numbers constructed on a basis complete details of which are perhaps not accessible to the investigator.

18. I have little doubt that the preceding paragraph could have been written with much more feeling and emphasis by anyone with more experience than I can claim to possess of the practical difficulties which attend the initiation of statistical projects in social or economic research. It is not too much to say that the acquisition of data for such projects is often an expertise in itself; and it requires no small effort of the imagination to conjecture how the statistical study of mortality on the one hand, and of economic behaviour on the other, would have fared if the respective availabilities had been reversed. To develop this fantasy we should have to envisage, among other strange phenomena, a system of national records concerned not with the births, marriages and deaths of individuals but with the details of their income, expenditure, consumption and so forth—mortality figures, on the other hand, being obtainable only as broad aggregates without distinction of age or sex save, perhaps, for an occasional sample inquiry on a more detailed basis. I am content, however, to leave the elaboration of the nightmare as an exercise for the reader.

19. Moreover, as regards the acquisition and handling of raw material there is another respect in which the economic statistician is at a disadvantage as compared not only with the actuary but also with the research worker in biology or other physical sciences. Just as the actuary can sort his records in a way which enables him to produce rates of mortality at a fixed age in relation to varying durations of assurance, so the laboratory worker can often arrange his experiment to exclude the effect of a particular variable and, by so doing, provide himself with a standard or 'control' which makes the significance of his subsequent discoveries much easier to assess. Such refinements of method are not usually possible in an economic study; the analytical process must be designed to produce results not from a basket of neatly labelled and graded eggs, but from an omelette concocted from an unknown recipe.

20. So much for the question of data. The other questions raised in paragraph 16 can perhaps best be approached by attempting to imagine some sort of mathematically designed Utopia in which all the factors exerting any

kind of influence on the consumption problem under investigation could be completely identified; in which the consumption itself, and the factors governing it, could be measured with absolute precision; and in which enough was known about the general nature of the relationship between the one and the other to enable that relationship to be expressed in terms of an algebraic formula connecting any one value of the dependent variable, i.e. any one member of the statistical time-series under analysis, with the corresponding values of the independent variables by means of a number of constants. In such circumstances, the number of sets of observations at our disposal being m and the number of constants in the *ex hypothesi* formula n ($<m$), the problem would not call for 'statistical' treatment in the generally accepted sense of that term; its elucidation would merely require the exercise of whatever mathematical ingenuity was necessary to solve any one set of n equations for the n constants. Every one of the $\binom{m}{n}$ sets would give exactly the same result.*

21. The whole theory and practice of statistical analysis has been evolved to circumvent the fact that the world in which we live does not conform to the specifications of a pipe-dreaming mathematician. Relationships between statistical series of the type we are considering have to be sought for in an atmosphere of imperfect data, relative probabilities and best approximations. The connexion between the dependent variable y and an independent variable x is not unique and immutable; the operative word is 'stochastic', and we must think in terms of regression lines and least squares solutions; indeed, the very distinction between the dependent and the independent variable becomes somewhat blurred when first we pass from algebra to statistics. This is the stage at which we ought to stop talking about variables and constants, and begin talking about variates and parameters; but I have preferred to retain the old-fashioned mathematical terms throughout.

* Although it serves the present purpose, I am not sure that this description of a mechanism exactly definable in mathematical terms is really complete. In our present state of knowledge, we find that most of the phenomena which we can measure are subject to change; is it not therefore possible to imagine that under conditions in which everything could be fully explained in terms of cause and effect there might be no such thing as a constant? If so, the algebraical formula expressing the effect in terms of the causes would consist solely of variables and the signs $+$, $-$, $\times$ and $\div$; in other words, such an expression as $x_2 = a_1 + a_2^{x_1}$ (the x's being 'variables' and the a's 'constants' in present nomenclature) might merely represent the best attempt we can yet make at, say,

$$x_2 = \left(x_3 x_4 + \frac{x_5}{x_6}\right) + (x_7 - x_8)^{x_1},$$

or something equally horrific; x_1 being the only factor which we can at present discern as having an influence upon x_2, and $x_3 \ldots x_8$ being factors which we have not so far thought of importing into our analysis but which, when blended in the manner indicated within the brackets, always produce (on the scale of accuracy to which we are working) the same result. There can, of course, be no question of 'solving' such an equation in the sense in which we would find a_1 and a_2 from $x_2 = a_1 + a_2^{x_1}$ with the aid of two or more sets of x_1, x_2; it would seem that both the structure of the right-hand side and the ingredients of which it is composed would have to be determined—if they ever could be determined—by trial and error on the basis of *a priori* reasoning. The whole notion seems rather fantastic; yet it cannot be denied that when we discover that something, hitherto regarded as invariable, is really the net result of variable factors we consider that we have added to our knowledge of the universe.

22. At this point I must refer directly to the account of principles and methods of multivariate analysis which I mentioned in paragraph 15. Four years ago Richard Stone read a paper to the Royal Statistical Society entitled *The Analysis of Market Demand* (*J.R.S.S.* Vol. CVIII, p. 287). In view of the title, it is safe to assume that this paper has not been widely studied in our own profession. From the illustration used in paragraph 16, it may already have been guessed that the paper deals with time-series of consumption—the consumption of a number of goods and services in the United Kingdom and the United States of America during the whole or part of the period between the two world wars. Stone describes his paper (paragraph 1.1) as 'a connected account of the problems involved in investigating market demand by means of multiple regression analysis', but although the paper naturally contains much that is special to this particular economic subject, it provides a review of lines of approach and technical methods of treatment which seems to me to be of wider interest and significance.

23. Stone points out that it is useless to start a piece of statistical work before the problem to be solved or the hypothesis to be tested has been clearly formulated. Before proceeding, therefore, to a discussion of statistical method, he advances a series of antecedent considerations which I have attempted to summarize (using very largely the phraseology of the original) as follows.

(1) If we are attempting to establish a stable relationship between a statistical series (representing the variable which for the particular purpose of our inquiry we choose to treat as 'dependent') and other statistical series representing 'independent' variables, we must try to bring in as far as possible all the important influences on the dependent variable.

(2) We must, somehow or other, decide upon the form of the relationship between the dependent variable and the independent variables—in other words, we must consider what the algebraic equation expressing this relationship is most likely to look like.

(3) We ought also—from our 'general knowledge', so to speak, of the subject-matter with which we are dealing—to specify such *a priori* expectations as we can about the signs and sizes of the constants with which the variables will be cemented together in the relationship equation.

24. In the discussion which accompanies this enunciation of general principles one point should be noted at this stage because it is specially relevant to the rough attempt which I have made in the Appendix to this paper to apply the principles to the analysis of non-economic data. In seeking a relationship between two or more time-series, to what extent should we introduce time itself as an independent variable? Clearly, if a statistical series exhibits a steady and consistent upward or downward trend from year to year, it is a relatively simple matter to express this trend as a mathematical function of the single independent variable, time; but, equally clearly, if we choose this easy way out, we cannot claim to have achieved much in the way of analysis to elucidate *a priori* expectations or to expose the nature and extent of the relationship between different influences on, or attributes of, the life of a community. Stone recognizes, however, that some of these influences or attributes may be hard to express in the form of statistical series; in such a case, the introduction of time as an independent variable is a permissible—indeed, a desirable—*faute de mieux* arrangement.

Thus the purpose of introducing time as a determining variate is simply to take care of sources of continuous systematic variation which have not been introduced explicitly and which, together, tend to push the dependent variate up or down through time. While time cannot be ignored, a large coefficient of time is disturbing in the sense that it indicates an important source or sources of systematic variation that cannot be measured and possibly cannot even be given a name. On the assumption that the general and specific economic factors are an effectively complete catalogue of economic sources of systematic variation, time will stand for systematic factors which are not economic but usually social in character.

25. As regards the second principle stated in paragraph 23 above, Stone suggests that, in the first place, the constants in the relationship equation should be capable of interpretation in the terms which we ordinarily employ in thinking or talking about the subject under investigation and that, where possible, simple forms should be given as a first approximation to these general concepts; and secondly—as a purely practical condition—that the determination of the values of the constants which give the best fit should not be unduly difficult and costly. Consideration of the second principle as amplified by these conditions brings us, I feel, to the crux of the whole problem of statistical analysis. It would appear at first sight that such an analysis can have little real purpose or meaning unless, in designing the lines on which it is to proceed, we give full weight to everything which our common sense and our general knowledge lead us to believe or to expect in regard to the nature of the relationship which we are seeking to define in mathematical terms. But what happens if, having done this, we find ourselves involved in a mathematical complex which is beyond the capacity of our technical equipment to resolve? In seeking an answer to this question we are, I think, forced to give further consideration to the real aim and object of the analysis itself. It will, however, be better not to pursue this rather vital matter further until we have tried to see to what extent the difficulty is likely to arise in practice.

THE MULTIPLE REGRESSION EQUATION

26. In economic concepts of the type with which Stone's paper is concerned, the question posed in the previous paragraph does not arise if it is assumed that the elasticity of the dependent variable y with respect to each of the independent variables x is constant $\left(\text{i.e. } \frac{\partial y}{y} = a\frac{\partial x}{x}\right)$, and that the time factor can suitably be expressed as a geometrical progression; for these conditions immediately suggest

$$y = \mathrm{A}x_1^a x_2^b x_3^c e^{rt}$$

as the form of the relationship equation in a case where (other than time, t) there are three independent variables x_1, x_2, x_3. If this equation is written in the form

$$\log y = \log \mathrm{A} + a \log x_1 + b \log x_2 + c \log x_3 + rt,$$

it is at once seen to resemble the standard equation of multiple regression analysis in which a dependent variable is expressed as the sum of a number of terms each of which is obtained by multiplying an independent variable into a function—the 'regression coefficient'—of the measures of correlation between the dependent variable and each of the independent variables. If there is more than one independent variable such an equation cannot be described as representing a regression *line*; in the case of two independent variables we can talk of a regression *plane*, but no physical analogies are available for any higher number unless we choose to talk of *hyperplanes* in n dimensions. In any mathematical sense this limitation is, of course, quite immaterial; what is material is the form of the right-hand side of the equation, any term of which is a function of one and only one of the independent variables, the links between the various terms consisting invariably of the sign of addition.* The form in which the independent variables are introduced need not be the same in each case, nor need it be the same form as that which embodies the dependent variable on the left-hand side of the equation. That is to say, if in a particular field of inquiry we knew enough about the nature of the relationship to believe, or were led by common sense or general *a priori* reasoning to expect, that y was connected with one independent variable linearly, with another logarithmically and with a third exponentially, we should begin our regression analysis by correlating y with x_1, $\log x_2$ and $\exp(x_3)$ in order to arrive eventually at a multiple regression equation of the form

$$y = ax_1 + b \log x_2 + ce^{x_3}. \tag{1}$$

The right-hand side of this equation does not begin with a detached constant because the usual starting-point of correlation technique is to standardize the data by expressing the statistical series representing the dependent and the independent variables (or, as in the case of x_2 and x_3 above, the predetermined functions of the variables) as deviations from their respective arithmetic means and adopting the respective standard deviations of the series as the units of measurement. Thus in a simple bivariate case the regression equation

$$y - \bar{y} = r \frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

can be written either as $$y = \left(\bar{y} - r\frac{\sigma_y}{\sigma_x}\bar{x}\right) + r\frac{\sigma_y}{\sigma_x}x$$

(the term in brackets being a constant)

or as $$\frac{y - \bar{y}}{\sigma_y} = r\frac{x - \bar{x}}{\sigma_x},$$

i.e. the standardized form referred to above.

27. In deriving equation (1), our prior knowledge or expectation has enabled us to use $\log x_2$ and $\exp(x_3)$ as independent variables instead of x_2 and x_3. What it comes to, in short, is that in attacking any problem of relationships between statistical series by the method of multiple regression analysis, we automatically commit ourselves to the fundamental assumption that the relationship can be mathematically expressed by

$$\phi(y) = af_1(x_1) + bf_2(x_2) + cf_3(x_3) + \ldots,$$

where the forms of ϕ, f_1, f_2, f_3, ..., are already known and statistical series corresponding to them can be derived from the data before we begin our

* It is this purely additive connexion between the terms which, as I understand it, gives rise to the use of terms such as 'linear' and 'plane' no matter how many variables are involved. Thus $\log y = \log x_1 + 2 \log x_2$ is describable as a linear regression equation or as representing a plane of regression although the corresponding expression $y = x_1 x_2^2$ connotes a curved surface.

attempt to assign values to the coefficients $a, b, c, \ldots$*. If we cannot determine these forms beforehand—that is to say, if our *a priori* knowledge or expectation amounts to no more than a general idea (substantiated, perhaps, by some correlation calculations) that y is associated with $x_1, x_2, x_3, \ldots$, *in some way or other*, then we can do no more than ring the changes on various combinations of forms for $\phi, f_1, f_2, f_3, \ldots$, beginning probably with simple hypotheses such as

$$y = ax_1 + bx_2 + cx_3 + \ldots,$$

or

$$\log y = a \log x_1 + b \log x_2 + c \log x_3 + \ldots,$$

in the hope that after a considerable amount of very dull arithmetic we may hit upon a set of assumptions which fit the data. Although the tool-kit of regression analysis contains a device for testing whether an assumption that the relation between two variables is linear is statistically valid, this test gives no indication whether the relation, if not linear, is logarithmic, parabolic, or anything else; and indeed, the whole analytical technique as so far developed cannot, it seems to me, provide any automatic indicator of this kind. We simply have to proceed by trial and error, exercising our common sense and ingenuity to the full in formulating basic assumptions the exploration of which is within the capacity of our data and our technique. Statistical analysis cannot perform the miracle of determining the mathematical pattern of the relationship which we seek—that must come out of our own heads. What the analysis may be able to do is to tell us whether the pattern we have chosen is a good one; but even this it can only do if we ourselves know what we mean by 'good' and if our algebra is competent to solve the equations which result from the application of our chosen criterion of 'goodness'.

EVALUATING THE CONSTANTS

28. This brings us to the operation of determining numerical values for the constants in the formula which common sense and general knowledge have led us to write down as an expression of the particular relationship we are studying—i.e. to what, in the two-dimensional case, we usually describe as curve-fitting. The two methods of curve-fitting most generally in use in actuarial or other bivariate work are the method of moments and the method of least squares. The former proceeds on the hypothesis that the results of a particular series of arithmetical operations on the data themselves should reproduce, or be reproduced by, those of the corresponding algebraic operations carried out on the relationship formula. The latter seeks the unique set of numbers which, when substituted for the symbols which denote the constants in the relationship formula, will make the sum of the squares of the differences between the observed and calculated values of the dependent

* We could obviously go a stage further. $f_1, f_2, \ldots$, need not be explicit functions in $x_1, x_2, \ldots$, only, if our *a priori* knowledge enables us to postulate a function such as $f(x_1, x_2)$ and write

$$\phi(y) = af_1(x_1, x_2) + bf_2(x_3) + \ldots,$$

in which case $f(x_1, x_2)$ becomes an independent variable in place of $f(x_1)$ and $f(x_2)$. But whether the f contains x_1 only, or x_1 in combination with other x's, we must form a series for it from our statistical data before we can begin to construct the regression equation. A combination of two or more series of varying degrees of reliability would, however, seem to give rise to difficult questions about the margin of error in an independent variable 'manufactured' in this fashion.

variable as small as possible. These two methods would presumably never have come into general use if they were not justifiable either on grounds of principle or on grounds of expediency; and it seems to me that the latter are much more easy to find than the former. In the case of actuarial work in general—by which I mean the 'graduation' of an age-distribution or a series of rates specific to age—I would suggest that questions of principle hardly arise at all, for the reason that they have been encountered and disposed of at an earlier stage of the investigation, i.e. during the process of pruning the heterogeneous elements in the data in the manner described earlier. If the net body of experience which survives this operation is at all substantial, all the actuary really needs is a systematic means of choosing between the various smooth curves, all adhering more or less closely to the observed values, which he could produce by freehand drawing. This is also true, in a very broad sense, of many of the frequency distributions with which the Pearsonian families of curves were originally designed to deal. The method of moments, if indeed it was not actually invented for the fitting of curves of this type, was at any rate found to be a very convenient device for the purpose; and the justification of the method rests, I suggest, on this quality of convenience rather than upon any supposition that the method itself satisfies any *a priori* criteria derived from general considerations of the nature of probability or anything of that kind. This point of view is not, I think, affected by the circumstance that for certain types of curve the results of using the method of moments are identical with those obtained from the method of least squares. It is interesting to note in this connexion that an alternative method of fitting a curve by the use of moments which has recently been put forward (see Sichel, *J.R.S.S.* Vol. cx, p. 337) is commended by its author not on any grounds of principle, but because it can produce equations which are easily soluble in certain cases where the ordinary method of moments produces equations which are either cumbrous or insoluble; and any addition to the apparatus of statistical analysis which can show practical advantages of this kind is, naturally, to be welcomed.

29. The problems of solving the equations which result from attempts at fitting formulae to observed data are discussed below; at the moment I will merely express a personal view that, if the choice of ways and means of determining the constants in a relationship equation lies between two methods which are equally convenient from the purely mathematical point of view, that one is to be preferred which is the more intelligible and less artificial from the non-mathematical point of view. My vote in a Gallup poll on this subject would be in favour of the method of least squares, as against the method of moments, simply because I find it easier to envisage what the application of this method is doing to the actual data. After all, what we are doing to the data is more important in the long run than how we are doing it. A subsidiary attraction of the least squares method is the elegance of the partial differentiation device for the formation of the normal equations—although this also may be a question of personal taste. In the present context, however, this discussion of relative merits is a digression; the method of moments can, so far as I am aware, be employed only in bivariate problems, and we are dealing here with the analysis of a time-series in which several variables may be involved.

30. Although the choice of a method for fitting a curve to mortality data may not usually present the actuary with any awkward questions of principle, he does not escape the more practical difficulties of the matter. It is one thing to translate the conventions we have chosen to adopt for 'fitting' purposes into

a set of equations; it is quite another thing to find a method of solving the equations when they are written down. For example, although we have been using the Makeham formula $y = \mathrm{A} + \mathrm{B}c^x$ for many years, we have never been able to devise a formal method of deriving the three constants from the equations which result from the application either of least squares or of moments procedure. The commonly accepted practice of determining A and B on the basis of trial values of c is not open to any really serious objection because the possible range of values for c is limited by the homogeneous character of the data; nevertheless, it would have been pleasant if our pre-occupation with this particular expression could have resulted in a contribution to pure algebra. As things are, if instead of using the trial value method we seek to find c by eliminating A and B from the curve-fitting equations, we are faced with an unwieldy polynomial to the reduction of which we can bring only the melancholy knowledge that no method exists for the formal solution of any equation of higher degree than the biquadratic. This takes us back again to 'trial and error' methods, and although the various text-book devices for the approximate solution of quantics are neat and ingenious, it cannot be denied that most of them are more than a little laborious.

31. The fact is that in the case of the expression $y = \mathrm{A} + \mathrm{B}c^x$ not only are the equations derived by the curve-fitting process not soluble for A, B and c by formal methods, but also there is no formal method of solving any set of three original equations

$$y_1 = \mathrm{A} + \mathrm{B}c^{x_1}, \quad y_2 = \mathrm{A} + \mathrm{B}c^{x_2}, \quad y_3 = \mathrm{A} + \mathrm{B}c^{x_3},$$

unless the three values of the argument x are connected by a very simple relationship. For if we write

$$x_2 = x_1 + \alpha, \quad x_3 = x_2 + \beta, \quad (y_3 - y_2)/(y_2 - y_1) = \mathrm{K}, \quad \text{and} \quad c^\alpha = z,$$

the equation obtained by eliminating A and B can be written

$$\mathrm{K} = \frac{c^\alpha (c^\beta - 1)}{c^\alpha - 1} = z \quad \text{if} \quad \beta = \alpha,$$

or $$z(z+1) \quad \text{if} \quad \beta = 2\alpha,$$

or $$z(z^2 + z + 1) \quad \text{if} \quad \beta = 3\alpha,$$

or $$z(z^3 + z^2 + z + 1) \quad \text{if} \quad \beta = 4\alpha,$$

with correspondingly simple equations if $\alpha = 2\beta$, 3β, or 4β; but outside this range the equation becomes a quantic amenable only to approximate methods of solution. If $c = e^\delta$, and $|\delta|$ is fairly small, a close approximation to δ is in fact provided by

$$\frac{2}{\alpha + \beta} \log \frac{\alpha \mathrm{K}}{\beta} = \delta - \frac{\alpha - \beta}{12} \delta^2.$$

So far as I am aware, there is no correspondingly simple approximate method of solving the three equations which result from the differentiation of the expression

$$\Sigma (\mathrm{A} + \mathrm{B}c^x - y)^2$$

with respect to A, B and c; but the fact that any set of three relationship equations is soluble for A, B and c by approximate methods suggests a line of approach which will be discussed further in a later paragraph in relation not only to this question of solubility but also to another problem of a different character, which I shall now attempt to discuss.

IDENTIFYING THE VARIABLES

32. An investigation of the relationship between a number of statistical series which can be handled by the ordinary methods of regression analysis does not encounter purely algebraic difficulties of this sort because, as has been shown in earlier paragraphs, such methods are based on a definition of relationship which is essentially 'linear' (using the word in the general sense described in the footnote to paragraph 26). The basic equation is invariably of the type

$$y = ax_1 + bx_2 + cx_3 + \dots,$$

$x_1, x_2, x_3, \dots$ representing either the 'independent' statistical series themselves or functions of those series which our *a priori* knowledge enables us to define and to calculate before we attempt to introduce them into the analysis of the series denoted by y. The application of least squares procedure to a system of equations of this type gives a similarly linear system for the determination of the constants a, b, c, ...; and any such system is soluble, either formally by determinants or other algebraic means, or by one or other of various methods of successive approximation. The values of a, b, c, ... so arrived at are the same as those obtained by the standard procedure of calculating multiple correlations between the series represented by y, x_1, x_2, x_3, ... The two processes are, in fact, demonstrably identical.

33. Now in seeking to establish a relationship between two or more statistical series by the method of least squares we are, in effect, setting up against each member of the series which we are treating as 'dependent' a calculated value which we consider to be more accurate, the difference between the observed value and the 'expected' value representing the extent of the 'error' in the former. This conception would obviously be altogether meaningless if it did not imply that we consider ourselves able to make a complete specification of the make-up of the dependent series and to compute its terms accurately by means of the various independent series. In other words, we assume not only that the members of the independent series are free from error, but also that the tale of them is complete. These are very big assumptions to make, but clearly we could not get very far without them. In the construction of a mortality table from a body of homogeneous data derived from life assurance records they are not, after all, so very unrealistic; but the position is different when we embark on statistical analysis in a field where the raw material is less reliable and we are unable to prune it with a view to discarding such sections as are specially likely to be affected by factors which we do not wish to include in our analysis. In such a case, while there may be little that we can do to improve the accuracy of the statistics themselves, it is evidently important that we should at least have some means of considering whether the attempt to define one series of statistics in terms of others is helped or hindered by changing our original specification.

34. As a simple illustration, let us consider a series of values of a quantity y which is being treated as dependent on a single variable x. A straight line is fitted to the whole series of observations by the method of least squares and it is found that the differences between the actual and the expected observations are very small, with two or three exceptions where they are conspicuously large. A formal significance test may indicate that the probability of chance deviations of this magnitude is very remote. How do we proceed? Do we congratulate

ourselves that we have discovered a very good working functional relationship, and take the view that by a freak of chance the particular set of observations on which our analysis is based contains not only one, but two or three examples of the 'odd case in a thousand' which does not conform? If so, we may as well discard the awkward cases altogether and derive from the remaining values a line which fits them even better. Or do we now bring into consideration a possibility which hitherto we have been content to disregard for the sake of simplicity—that the behaviour of y may be affected by some other independent variable (z, say) in addition to x? In practice the answer to this must depend on whether we have any means of representing z by a statistical series which can be imported into the calculations. Assuming that we have this means and that the relationship we postulate between y and z, as between y and x, is linear, we can transfer our analysis from the second to the third dimension and seek, not a straight line, but the plane which will give the 'best' representation of the compound relationship between y on the one hand and x and z on the other. If the introduction of z provides a useful contribution to the analysis, it is to be expected that such a plane will lie more evenly among the various points (x, y, z) than the straight line did among the various points (x, y), and that the sharp distinction between the many 'good' points and the few 'bad' points will tend to disappear. But how do we assess the straight line and the plane in regard to these qualities of goodness or badness or evenness? We have already selected the best that is to be had in both the dimensions we have visited so far; but for all we know there may be yet another source of variation which, if we could but take cognisance of it, might yield something better than either.

35. In the realm of economic statistics this difficult question of investigating the effects of making additions to (or subtractions from) the list of independent variables included in the relationship equation has been the subject of study for some time; the names of Frisch, Koopmans and Hotelling are among those particularly associated with it. In *The Analysis of Market Research* Stone gives an account of the method of confluence analysis devised by Frisch, and applies it to the investigation of the consumption problems with which the paper is concerned. The essence of the method is to make a comparative study of the effects of basing the relationship hypothesis on any one or more members of the stock of potential independent variables, the whole process being worked out through the medium of the multiple correlation functions employed in standard multivariate regression technique and translated into a series of diagrams known as bunch-maps. The amount of arithmetic involved is very considerable as, indeed, it is bound to be in any kind of statistical analysis which seeks to make the fullest use of all the raw material available. There is, however, no point in describing the method more fully here because, being built up on the basis of ordinary regression technique, it does not appear to be applicable to the study of relationships which cannot be expressed in a form amenable to treatment by that technique; and it is with awkward cases of this kind that the considerations discussed in previous paragraphs have involved us. What can we do about it?

SOME TENTATIVE SUGGESTIONS

36. Now, whatever may have been said in previous paragraphs about probing into the unknown, it is only fair to our own common sense to assume that we shall not embark upon the analysis of a statistical series without

knowing quite a lot about the subject to which the series relates. In particular, we shall be reasonably sure which are the major factors governing the behaviour of the series—i.e. which are the most important independent variables—and in what sort of mathematical shape they should be introduced into the relationship equation. The potential independent variables which it may or may not be beneficial to include in the analysis will, relatively speaking, be trimmings rather than essential basic ingredients. This suggests that our first rough approach to the relationship equation might be an expression involving only the major independent variables, the minor ones being regarded for the time being as hidden in the constants by which the variable terms in the equation are connected. Having got so far, our further procedure would naturally be to determine the constants by the method of least squares and to calculate an 'expected' value for each term of the dependent series. If, on this first approach, we have been successful in our choice of the major independent variables, and have adopted the right sort of mathematical pattern in introducing them into the relationship equation, the 'expected' series will compare reasonably well with the actual series; indeed, the comparison may be so good that we may decide not to pursue the search for further sources of variation. If, however, we want to get a better fit we shall inspect the A–E column more closely, noticing the run of the positive and negative deviations and their relative sizes. We are dealing with a time-series, and it may be that the run of the differences from year to year will, by reminding us of the behaviour of some other time-series, or in some other way, give us a clue to a further independent variable. It does not follow, however, that they will give us any clue to the mathematical form in which the new variable should be introduced into the relationship equation.

37. If our sole objective was closeness of fit regardless of principles, the easy and obvious course would be to treat our residual A–E as a linear correlative of the new variable—assuming that a statistical series representing the latter is at our disposal—and we would thus obtain a simple additional function to complete our relationship equation. But this course will not satisfy us if our attitude to the whole question of statistical analysis is based on the view that common sense, general knowledge and *a priori* expectations must come before goodness of fit and that no violence must be done to principles merely for the sake of easing technique. In such a case it would seem that, having discovered what appears to be a new variable, we must consider what sort of functional relationship may most reasonably be expected to subsist between it and the dependent variable, then modify our relationship equation accordingly, and start all over again.

38. There are, however, three objections to the procedure described in the last two paragraphs. The fact that it involves a lot of work is, perhaps, not in itself a very valid objection, because it must apply equally to any alternative method; but it does seem rather ridiculous that each stage of the approach to the final form of the relationship equation should involve the whole gamut of operations right down to the computation of the 'expected' values. The second objection—perhaps a somewhat metaphysical one—is that the use of least squares in these exploratory stages is a rather bizarre adaptation of a principle which presupposes that the specification of the 'expected' values of the independent variable is already as complete and accurate as we can make it. The third, and really weighty, objection is that unless the form of our relationship equation is very simple (in which case it might be amenable to ordinary regression technique and the whole problem would not arise) the

least squares equations defining the constants will almost certainly be insoluble by formal methods and probably very difficult to handle by approximate methods.

39. If the first approximation to the relationship equation (i.e. the preliminary form into which only the major or more obvious independent variables are introduced) is amenable to formal algebra or to approximate methods, we can overcome these objections to some extent. Suppose that there are n sets of observations and that the approximate relationship equation contains r constants; we can then form $\binom{n}{r}$ groups of equations, each of which will yield a unique set of values for each of the constants. Now the constants fall into two classes—those which are unlikely to be affected to any substantial extent by later modifications of the relationship equation, and those which have been introduced into the equation at this stage as temporary substitutes for functions of subsidiary independent variables. The $\binom{n}{r}$ values obtained for members of the latter class by the unique solution process cannot be expected to arrange themselves in any recognizable probability distribution, and there is, in fact, no particular point in computing them at all. As regards the constants of the more genuine type, however, the position is different. These are associated, *ex hypothesi*, with the more stable and outstanding characteristics of the series under analysis and it might be expected that, notwithstanding the disturbing influences exerted by the minor variables, the $\binom{n}{r}$ values would form a distribution with a recognizable mode which would indicate what we may perhaps call 'the most likely value in (or ? despite) all the circumstances'. We should also expect the values at the extremes of the distribution to be those derived from sets of equations comprising the terms of the statistical series most markedly affected by the subsidiary variables, and (as has already been suggested in paragraph 36) this might help to put us on the track of these subsidiary variables. Going a stage further, it would seem reasonable to take the view that of two alternative forms of the relationship equation that one is to be preferred which produces the more compact distribution of values of the constants around their modes.

40. There would, however, be awkward practical difficulties in the selection of the modal values of the 'genuine' constants if the number of such constants were at all considerable. Suppose there were K of them, $a_1, a_2, \ldots, a_K$. The values we require are not necessarily those given by the modes of the distributions of the $\binom{n}{r}$ values of a_1, the $\binom{n}{r}$ values of a_2, and so on, considered separately, but the mode of the $\binom{n}{r}$ members of the family $a_1, a_2, \ldots, a_K$ considered as a single unit. If the family had only two members, a_1 and a_2, we could plot one against the other on a diagram and look for the densest cluster of points, although it would, of course, be preferable to use some formal technique for the assessment of density. If there were three members a_1, a_2, a_3, the visual approach would require the construction of a wire model; and if $K > 3$, a formal calculus for identifying the densest cluster would be not merely desirable but essential.

41. It is interesting to speculate how values obtained for the 'genuine'

constants by this method would compare with those obtained by least squares. As the method has been suggested for use in cases where the least squares equations are insoluble, the question is rather academic; but it would, of course, be possible to compare results in a case where the least squares equations presented no difficulties. It may, indeed, be possible to answer the question from purely theoretical considerations; but although I have used the 'modal value' approach in the simple exercise worked out in the Appendix I cannot pretend to have given it much thought from the point of view of principle. There is, however, a practical point which ought to be mentioned, and that is the labour of solving the $\binom{n}{r}$ sets of equations.

42. If there were twenty terms in the statistical series under analysis, and the experimental relationship equation contained three constants, two of which were 'genuine' in the sense indicated earlier, we should be faced with the solution of 1140 sets of three equations although, as has already been pointed out, we need not bother to extract values for the non-genuine constant. This is a very simple example; with thirty terms and five constants there would be 142,506 sets of equations—moreover, the labour of solving each five-set would be much greater than the labour of solving each three-set. This suggests that the proposed method is altogether impracticable if the statistical series is at all long and the postulated relationship equation at all complicated, unless we can either turn the business of solving the sets of equations into a very rapid mass-production process or reduce, somehow or other, the number of sets of equations to be solved. The former course would probably involve the invention of new methods of approximate solution and the construction of some special type of machine for carrying them out, while the only method of reducing the number of sets of equations which does not seem to be fundamentally objectionable at first sight is by sampling. This may seem rather a quaint suggestion to make; but after all, we have brought sampling methods to a high level of efficiency and use them with confidence to an ever-increasing extent in the procurement of data. Why should we not have the courage of our convictions and introduce the sampling idea into the analytical processes themselves? I make no attempt to answer this question; I will simply draw attention to one point which would have to be covered in any attempt to reduce the computing work in this way. In the $\binom{n}{r}$ sets of equations each term of the statistical series occurs the same number of times as any other, viz. $\binom{n-1}{r-1}$ times; and this condition would have to be preserved in the sample if we wished to give equal weight to each term of our series. This would seem to suggest a need for some kind of stratification.

CONCLUSION

43. It is time to bring this rambling discussion to a close. In the earlier paragraphs of this paper I have tried to show that actuarial methods have been developed in conditions more favourable than those which govern the activities of the statistical analyst in other fields. It would be strange if habits of thought engendered by this difference in environment were not reflected to some extent in a paper written by any actuary who, having been rash enough to attempt a journey into unfamiliar territory, finds himself faced with awkward obstacles

of a kind not often encountered on the comparatively safe and well-trodden path of life contingencies. It is possible that the dilemma which I have described in paragraph 25 as 'the crux of the whole problem of statistical analysis' only appears as a dilemma when seen through actuarial spectacles; we have achieved such a smooth and steadfast working relationship between mathematical theory and practical technique that it is not altogether easy to accept the idea that useful results can be obtained in circumstances in which principles may have to be a little elastic if methods are to be found to apply them. But the degree of elasticity must, I suggest, depend on the purposes to which the analysis is to be put.

44. If we are making a statistical attack on a subject of which we have little general knowledge, obviously the common-sense practical course is to open up the problem in a preliminary way by adopting methods which present no difficulties on their own account. There is no question here of giving the correct mathematical expression to our *a priori* considerations, because we have none; and, after all, this is the very type of situation for which the whole technique of correlation and regression was evolved in the first instance. But if we know enough of the subject under analysis to be able to postulate with some confidence the form which the relationship equation really ought to take, to what extent are we justified in modifying that form in order to bring it within the scope of our technique? The answer here, I think, must depend primarily on what our object in undertaking the analysis really is.

45. It may be that the general theory of the subject under analysis is so well established that there is no doubt whatever either about the form of the relationship equation or about the identity of the independent variables included in it. In that event, the object of our analysis is most likely to be to discover the relative magnitudes of the various constants in the equation, since this knowledge will enable us to put our general doctrine on an even more exact footing. In such circumstances it seems to me that we should not tamper with the relationship equation in any way, but should address ourselves to the problem of finding a technique which will deal with it, even if only in an untidy and imperfect fashion. To adapt principles instead of methods would simply be to bury our heads in the sand.

46. We may, however, be less interested in the mathematical pattern of the relationship for its own sake than as a means of estimating missing terms in the statistical series under analysis. Here it is very necessary to make a distinction between interpolation and extrapolation. Within the range of known values there may be little to choose between the results obtained by the use of formulae of radically different types; but if we venture any considerable distance outside the range we run into dangers of the sort which we can all remember being warned against in student days. It would seem unlikely that these dangers become any less when we pass from bivariate to multivariate distributions. The tabulated example on page 203 is sufficient to show how important is the choice of formula for extrapolation purposes.

47. Goodness of fit, then, is no sort of criterion for forecasting purposes; we must be confident that the relationship equation is of the right pattern. The use of multivariate analysis as a means of forecasting is obviously also limited by the extent to which the future course of the statistical series used as independent variables can be prognosticated; and in many cases this must mean that we can only say how the dependent variable is likely to behave on the basis of stated assumptions about the behaviour of the independent variables. At

Independent variable (t)	Dependent variable (u)				
	Actual values	Expected values by formula			
		(1)	(2)	(3)	(4)
−5	—	274	297	304	352
−2	222	220	222	223	224
−1	203	202	201	201	200
0	178	184	182	181	181
+1	167	166	165	165	164
+2	150	148	150	151	151
+7	—	58	92	118	107
+10	—	4	68	130	91

(1) $u=a+bt$. (3) $u=a+bt+ct^2$.

(2) $\log u=a+bt$. (4) $u=\dfrac{a}{b+t}$.

first sight it might appear that this is a minor drawback in any analysis in which time itself is a major factor; but, as was pointed out in an earlier paragraph, the introduction of a purely secular variable may be no more than a makeshift device necessitated by our inability to specify, and to define in terms of statistical series, all the influences which have in the past combined to produce a steady secular trend. And this is a reminder of the fundamental limitation on all forecasting by formula—a limitation so patent that it is easily overlooked; in a changing world no guarantee can possibly be given that a pattern of relationships which has prevailed in the past will persist without substantial modification in the future.

48. To what extent are the capabilities and limitations of statistical analysis of interest to the actuary? There are, I suggest, several answers to this question. The scope of actuarial activity is not limited in any way by the historical circumstances described at the beginning of this paper; as time goes on, an increasing proportion of the work of the profession tends to be located in fields where data are less complete and reliable and where methods of dealing with them less firmly established. The actuary ought therefore to be fully competent to handle any tools that have been designed for use in conditions of this kind; and, what is more, he should be prepared to offer any contribution which his own training and experience may enable him to make to the development and improvement of existing methods. I have tried to show that the greater our fund of qualitative knowledge of a given subject the more difficult it may sometimes be to translate our ideas into the language of quantities; so that as knowledge increases so must our skill as applied mathematicians and statistical analysts be continually heightened and refined to keep pace with it. We cannot for ever content ourselves with polishing and repolishing the ideas and the processes which have served us so long and faithfully in the well-charted field of life contingencies, without incurring some eventual risk of intellectual stagnation. And within the field of life contingencies there are at least two subjects—sickness and fertility—which would seem to invite more systematic study from the point of view of their association with general social and economic factors than we may find it necessary to undertake when our primary concern is with rates of benefit and contributions in connexion with an insurance system. Lastly, there is an intrinsic interest—intellectual, aesthetic, artistic or what you will—in multivariate analysis. The trouble is that it

involves so much computing and so—unlike the Third Programme or The Times Crossword—we cannot get very far with it in an armchair.

49. It is customary for the author of an Institute paper to express the hope that it will stimulate discussion. I certainly present this paper in expectation of a good deal of criticism; the problems which it attempts to discuss have for me a fascination which far outruns my knowledge of the current literature of statistical method, and it may well be that I have found difficulties where none exist and stumbled against obstacles which have already been surmounted or by-passed. But I would like to say one word in advance of any suggestion that the paper reveals an excessive preoccupation with mathematical forms. The principles and methods which we use with such ease and confidence in our daily work derive to a very large extent from the pioneer activities of men who were eminent mathematicians in their own right; and it would seem a poor tribute to their endeavours to imagine that the capacity and vigour of the profession can be fully maintained unless we continue to take full advantage of everything that mathematics has to offer.

50. I am greatly indebted to Mr Stone, whose work on economic analysis did much to inspire the train of thought which has ultimately sought expression in this paper. Mr Stone was good enough to read the paper in draft at very short notice and to write to me about it. I am grateful also to Dr Glasspoole of the Air Ministry for putting me on the track of the meteorological data used in the Appendix; to the Royal Meteorological Society for permission to reproduce these data from the Society's *Journal*; to the Controller of His Majesty's Stationery Office for similar permission regarding the official mortality figures on which the calculations in the Appendix are based; and to numerous colleagues—both senior and junior—in the Government Actuary's Department for helpful suggestions, for readiness at all times to participate in the discussion of ideas, and for some valuable assistance in experimental computing.

APPENDIX

MOVEMENTS IN THE NATIONAL DEATH-RATE BETWEEN THE TWO WARS

51. The object of the calculations which follow is to illustrate in a simple manner the foregoing discussion of principles and methods. They do not purport to give a detailed analysis of mortality trends and fluctuations during the period selected for examination. For a full study it would be necessary, *inter alia*, to take the seasonal rhythm of the death-rate into account; this would mean using rates calculated on a quarterly or even a monthly basis, thus making the whole matter more complicated than is necessary for the present purely illustrative purpose. With simplicity again in view, the function chosen for study is the standardized death-rate for persons, so that the effects of changes in sex- and age-distribution do not come into the picture. The standardized death-rate for persons in England and Wales is, of course, merely an average of the death-rates in age-groups for each sex, arrived at by using a set of fixed weights derived from the 1901 population. The rates used in the present study are taken from tables 5 and 9 of Part I (Medical) of the Registrar-General's *Statistical Review* for 1941. Since the publication of this volume the Registrar-General has adopted a new method of comparing the mortality of one year with that of another, on the ground that the age-distribution of the

population has changed so considerably since the beginning of the century that an index based on the old system of weights has become unrealistic. This circumstance, however, merely serves to emphasize the academic character of the present study: it does not invalidate it statistically in any way.

THE VARIABLES

52. The most casual glance at any time-series of twentieth-century death-rates is enough to indicate that from the point of view of factor analysis the dominant variables are secular in character. This feature is so generally recognized that it is unnecessary to illustrate it with a diagram; the extent of the decline in the standardized rate is revealed by Table 3. Superimposed upon the general trend are annual fluctuations which, when the series is charted, produce an irregular saw-tooth pattern; but with one notable exception—the influenza epidemic of 1918—they are not sufficiently marked to distract attention from the broad downward sweep of the underlying curve. As the Registrar-General has remarked on more than one occasion in the *Annual Review*, the main factors which produce these fluctuations from year to year are the incidence of epidemics and the meteorological conditions experienced during the winter months.

53. Here, then, are three factors by reference to which an analysis of the time-series of standardized rates can be undertaken—time, epidemics and winter weather. As regards the first of these, we may recall what was said in paragraph 24 of the Paper about the use of time itself as an independent variable. The social and economic influences which have combined to produce the steady reduction in mortality during the past half-century are numerous and complex. To take specific account of each separate factor we should need statistical series relating to standards of housing, sanitation, nutrition, personal hygiene, real wages and probably many others, as well as a measure of the rate of advance in medical science and treatment; and if such series were available their introduction into the analysis would be complicated by the need to take account of time-lag effects which would probably vary considerably between one series and another. Thus at the very outset of our inquiry we are faced with substantial problems in the procurement of data. But the concern of this Appendix is with analytical processes, not with a survey of possible sources of statistical material; so that although we are not merely rounding up residual influences, but dealing with a group of major variables, the practical course is to treat the various possible specific factors as a conglomerate epitomized by time itself.

54. Having started on the path of simplifying our working methods for the sake of convenience, we may as well take the further step of deciding to represent the downward secular trend by a smooth continuous 'concave-up' curve without maxima, minima, or points of inflexion, such as is given by a declining geometric progression. In real life, things do not happen quite in this way; progress may be continuous, but from time to time its pace slackens or quickens. To depict these changes in tempo would mean introducing an element of waviness into our curve by adopting a formula capable of giving points of inflexion at the appropriate stages in the time-series. We could not, however, locate these stages without some further study of the individual factors which we have already decided to represent summarily by a single time-factor. In the period between the two world wars which is covered by this analysis there does not appear to have been any single social, economic or medical development

with a conspicuous bearing on the course of the mortality rate; but if the analysis were extended into the current decade it would be necessary to take cognizance of the general adoption in medical practice of penicillin and the sulphonamides.

55. The next variable on our list is epidemic mortality. Public hygiene in this country has now reached a stage at which, so far as adult mortality is concerned, the only important epidemic is influenza. Since the beginning of the century there has been an enormous drop in mortality from the four main epidemics of childhood—scarlet fever, whooping cough, diphtheria and measles. The death-rate from these diseases does not seem to be capable of such violent fluctuations from year to year as the death-rate from influenza, and in the present context we are interested in epidemic mortality only from the point of view of its variability. It is no part of our purpose to attempt to investigate the periodicity of epidemics, nor can we subdivide the standardized mortality rates for any of these diseases between the endemic and the true epidemic except on some arbitrary basis which would detract from the factual value of our raw material. The easy way of dealing with the epidemic variable is therefore to confine attention to influenza, simply deducting the whole of the standardized influenza death-rate from the standardized death-rate for all causes. This course conveniently frees us from any obligation to investigate the possibility of a connexion between influenza epidemics and meteorological conditions. It also ignores the possibility that in a bad influenza year the number of deaths originating from an attack of influenza may be rather greater than the number classified under that heading for the purpose of the Registrar-General's standardized rate; however carefully the causes of death may be certified it seems almost inevitable that some deaths will be attributed to sequelae which, if all the facts were known, would come under the influenza heading. In table 9 of the Registrar-General's 1941 volume previously cited, there does seem to be some association between the rates for influenza and those for pneumonia during the period 1921–30; there were four bad influenza years in this decade, and on each occasion the pneumonia death-rate shows a divergence from the general trend. But here again any attempt to adjust the figures would import an arbitrary element into the analysis, especially as three of the four years in question had a low average winter temperature.

56. It is clear that any investigation of the trend of mortality by reference to recorded causes of death may be attended by some statistical risks if the period covered by the investigation is at all lengthy. Fashions in diagnosis may change; the technique of completing the death certificate may also change; and from time to time changes are made in the system of statistical classification. In the best of circumstances, these considerations must have some bearing on the homogeneity of a long series of disease-specific death-rates. In the present instance, causation comes into the analysis only to the extent that the influenza death-rate has been selected to denote the epidemic variable, and during the period 1901–40 this death-rate averaged only $2\frac{1}{2}\%$ of the rate from all causes. Even so, it seemed best to limit the span of the inquiry to the period 1921–38,* thus evading any possibility that the pandemic of 1918 may have induced a change in 'influenza-consciousness' and so rendered the pre-

* The war periods themselves were excluded for obvious reasons. The standardized rates for the years 1915–20 and 1940–41 are based on civilian mortality only and are radically affected by the withdrawal of healthy lives from the civilian population for service with the Forces.

and post-1918 series of rates not completely comparable. Even within this short period, it so happens that, as a result of a change in the system of classification, the disease-specific death-rates for years prior to 1931 given in the Registrar-General's table 9, already referred to, are not comparable as they stand with those for 1931 and later years; but conversion ratios are appended to the table by means of which the older rates can be adjusted, and the adjustment has of course been made in the figures used in this analysis, although I must confess that I do not quite understand how changes in classification can be implemented by the use of a constant factor.

57. On the question of epidemics it only remains to add that during the last forty years there have been two or three outbreaks of diarrhoea and enteritis (mainly infantile) on a scale which has appreciably affected the standardized over-all death-rate. It was discovered during the course of the analysis that one of these outbreaks occurred at the beginning of the period under review, but no change was made in the coverage of the epidemic variable on this account. A further reference will be made to the point in a later paragraph.

58. We come now to the meteorological variable. The association between sickness and mortality on the one hand, and weather conditions on the other, is a subject on which much could be written. It is, indeed, to most people a subject of perennial interest. We are as fond of discussing the weather as we are of discussing our ailments, and more often than not we find ourselves speaking of both in the same breath. But in the present context we are not concerned to argue the hypothesis that a severe winter is likely to give rise to more ill-health, and a higher mortality rate, than a winter which is less severe; our immediate purpose is to find some means of introducing this relationship into a statistical analysis of a mortality time-series.

59. The meteorological character of two winters can be compared on a number of different bases according to whether we are thinking of temperature, rainfall, snowfall, humidity, fog, sunshine, wind or changeability in general. For some of these elements statistical time-series are available; but as far as I know there is no general meteorological index combining them in the way that a price index, for example, is obtained by calculating the weighted average of the prices of a number of individual goods or services. Indeed, it would seem that the content and the method of construction of any such composite index must depend on the context in which it was to be used; an index for use in the statistical analysis of morbidity might be of little use to the student of traffic movement or agricultural productivity. In making the calculations in this Appendix, I have made the assumption that, so far as mortality is concerned, a major relevant meteorological factor is the presence or absence of 'cold spells', and that as really wintry conditions do not often inflict themselves upon us until after Christmas, the character of any year in this particular respect can be roughly represented by the mean air temperature of the three months January–March. My statistical series for the meteorological variable has therefore been obtained by averaging the figures given for each of these months by Glasspoole and Hogg in *Serial Monthly Values of Mean Temperature over the British Isles*, 1881–1940, *and Annual Values*, 1866–1940, *Journal of the Royal Meteorological Society*, Vol. LXVIII (1942) Table 1A (England and Wales). These temperatures are the means of the sea-level equivalents of the daily maximum and minimum readings taken at observation points widely distributed over the country. Clearly they are not ideal for our purpose; the entire population does not reside at sea-level

nor—unless by a grand coincidence—is it geographically distributed in a manner consistent, from the point of view of averaging, with the spread of the temperature observations. But the rough-and-ready character of the calculations in general affords some justification for dismissing these considerations as *de minimis*.

THE DATA

60. The data described in the foregoing paragraphs are assembled in Table 1, which shows, for each of the calendar years 1921–38, the standardized death-rate per thousand persons in England and Wales from all causes, from influenza, and from all causes other than influenza, and the mean daily temperature of the first quarter of each year in degrees Fahrenheit; the four series being denoted by m, f, m' and T respectively. That the subtraction of f has some smoothing effect is apparent from the table; the association of m' and T is better indicated by Diagram 1, in which the two series have been plotted (the T-series upside down) against time; most of the peaks and valleys in the one case are seen to correspond to peaks and valleys in the other.

Table 1. (See paragraph 60)

Calendar year	m	f	m'	T
1921	11·27	·19	11·08	44·5
1922	11·59	·46	11·13	40·9
1923	10·25	·17	10·08	43·6
1924	10·66	·38	10·28	40·3
1925	10·62	·25	10·37	42·3
1926	10·03	·18	9·85	43·6
1927	10·53	·43	10·10	42·7
1928	9·85	·15	9·70	43·3
1929	11·37	·53	10·84	38·3
1930	9·49	·09	9·40	41·5
1931	10·08	·26	9·82	40·1
1932	9·73	·23	9·50	41·9
1933	9·78	·40	9·38	41·7
1934	9·29	·10	9·19	40·9
1935	9·01	·13	8·88	43·1
1936	9·19	·11	9·08	41·0
1937	9·24	·30	8·94	41·5
1938	8·52	·08	8·44	44·7

THE RELATIONSHIP EQUATION

61. Before attempting to choose the most suitable form of relationship equation in the light of such *a priori* considerations as we can muster, let us see what sort of result we get by means of two formulae which present no procedural difficulties in fitting. If we write

$$m' = a + bt + c\mathrm{T}, \tag{1}$$

or

$$m' = ae^{bt}\mathrm{T}^{c}, \tag{2}$$

i.e.

$$\log m' = \log a + bt + c \log \mathrm{T},$$

t representing time, i.e. the calendar year of observation *minus* whatever calendar year is chosen as origin, we can evaluate the constants for both equations either by calculating the appropriate correlation coefficients or by means of a linear system of equations found by the straightforward application of the method of least squares. The results are shown in Table 2. So far as fit is concerned, there is not much to choose between the two formulae. Each produces bad results for the years 1921, 1923, 1924, 1929 and 1930. As regards the other 13 years, A–E is less than 1% of A in eight cases under formula (1) and in nine cases under formula (2); in the remaining years the error is, generally speaking, between 1 and 2%.

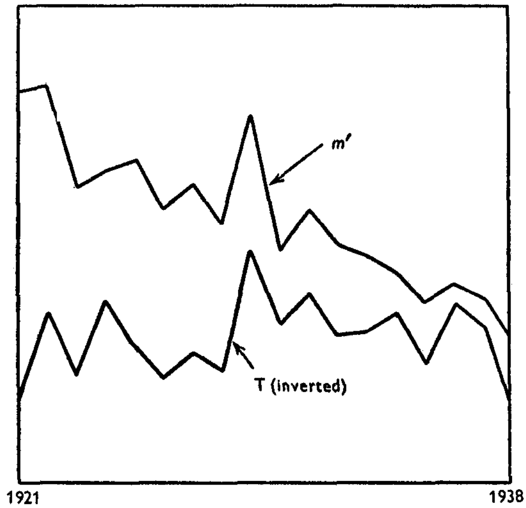

Diagram 1.

62. From the point of view of principle, formula (1) can of course be shot to pieces immediately. Whatever the object of our analysis, we should find it difficult to justify the use of a formula which carries the implication that about three-quarters of a century hence the rate of mortality (even a standardized rate which by then may be very artificial on other counts) will become negative. Under formula (2), m' will at least be always positive; the formula implies that by the end of the century it will have fallen to rather less than half its 1938 value, and no one can say that there is anything inherently ridiculous in contemplating such an eventuality. The formula gives a value for m' at the beginning of the century which is nearly 15% below the actual figure; this difference is considerable, but of no real importance because, having perfectly good actual values over a long series of past years, we stand in no need of backward extrapolations.

63. The real objection to formula (2) is that it commits us to a hypothesis which, in the present state of our knowledge, we are not justified in adopting, viz. that there is no limit to the extent to which the rate of mortality can continue to fall. This hypothesis would seem to involve the assumption of a continuous and indefinite extension of the span of human life, for which the evidence is so far lacking. Until such evidence is forthcoming it appears more realistic to consider the secular improvement in mortality as something which is altering the shape of the curve of deaths without lengthening the axis on which the curve is based, and to have regard to the distinction between 'senescent' and 'anticipated' deaths drawn by R. D. Clarke in his Centenary

Table 2. (See paragraph 61)

Calendar year	t	Expected m' by formula		A–E by formula	
		(1)	(2)	(1)	(2)
1921	−9	10·55	10·55	+·53	+·53
2	−8	10·91	10·96	+·22	+·17
3	−7	10·41	10·40	−·33	−·32
4	−6	10·73	10·77	−·45	−·49
5	−5	10·32	10·31	+·05	+·06
6	−4	10·02	9·99	−·17	−·14
7	−3	10·01	9·98	+·09	+·12
8	−2	9·80	9·76	−·10	−·06
9	−1	10·34	10·38	+·50	+·46
1930	0	9·78	9·75	−·38	−·35
1	+1	9·84	9·82	−·02	—
2	+2	9·46	9·44	+·04	+·06
3	+3	9·36	9·34	+·02	+·04
4	+4	9·33	9·32	−·14	−·13
5	+5	8·91	8·90	−·03	−·02
6	+6	9·06	9·06	+·02	+·02
7	+7	8·86	8·87	+·08	+·07
1938	+8	8·30	8·36	+·14	+·08
Σ (A–E)				+1·69 −1·62	+1·61 −1·51
Σ (A–E)^2				1·13	1·06

(1) $a=15{\cdot}38$, $b=-{\cdot}131$, $c=-{\cdot}135$.
(2) $a=95{\cdot}57$, $b=-{\cdot}0135$, $c=-{\cdot}613$.

(*Note*. The fact that Σ (A–E)≠0 under formula (1) is due to the drastic cutting down of decimal places throughout the calculations.)

Assembly paper *A Bio-Actuarial Approach to Forecasting Rates of Mortality*. This means that in our relationship equation we must remove some part of m' from the influence of the time-factor. We have already decided (paragraph 54) to assume that secular influences on mortality may, for the purposes of the present analysis, be represented in a broad general way by an exponential type of function, so that, but for the introduction of the meteorological variable, we should be able to write, for our relationship equation,

$$m' = \mathrm{A} + \mathrm{B}e^{-\delta t},$$

that is to say, a time-version of the old familiar Makeham age-formula with $c <$ instead of > 1.

64. The annual fluctuations in m' which for present purposes we are attributing to T are too considerable to enable us to use the values of m' for the individual years 1921–38 as a means of exploring the time-relationship independently of the secondary variable; but if there is any stability and continuity about this time-relationship it is surely permissible to seek information about A, B and δ in a wider context, particularly if this helps us to escape from the disturbing influence of T. Now almost the only statement we can make with any confidence about weather conditions in this country is that they are likely to vary considerably from year to year, so that we should expect that even over short spans of years annual fluctuations in T would tend to cancel each other out. Table 3 broadly confirms this. The values of m and f in this

table are the standardized rates for five-yearly periods as given in the 1941 *Annual Review*; they are, of course, almost identical with those obtained by averaging the figures for individual years. The value of m for 1936–40 has been roughly adjusted to exclude the air-raid mortality of 1940. The averages in the last column are not quite as flat as we should like to see them; but no value differs from the over-all mean (41·4) by as much as a degree, the average deviation irrespective of sign being only half a degree.

Table 3. (See paragraph 64)

Calendar years	t	m	f	m'	T
1901–05	1	15·96	·17	15·79	41·1
1906–10	2	14·44	·21	14·23	40·5
1911–15	3	13·69	·15	13·54	41·9
1916–20	4	13·43	·88	12·55	40·6
1921–25	5	10·87	·29	10·58	42·3
1926–30	6	10·26	·27	9·99	41·9
1931–35	7	9·58	·22	9·36	41·5
1936–40	8	8·95	·16	8·79	41·5

65. The next step is to determine A, B and δ by the method proposed in paragraph 39, using the quadratic in δ given in paragraph 31. In these calculations the group value of m' for 1916–20 has been excluded for the reason stated in the footnote to paragraph 56; the same consideration applies to the group value for 1936–40 so far as the mortality from September 1939 is concerned, but the period involved is short and any attempt at adjustment would be entirely arbitrary.

66. The thirty-five values of A, B and δ derived from the $\binom{7}{3}$ sets of equations provided by the second and fifth columns of Table 3 are arranged in Table 4 in ascending order of the δ values. There is no difficulty in selecting as the densest cluster the four sets of values between the lines across the middle of the table. Averaging these, we obtain

$$m' = 5{\cdot}66 + 11{\cdot}99e^{-{\cdot}168t}$$

or, on conversion to the time-scale of Table 2,

$$m' = 5{\cdot}66 + 4{\cdot}09e^{-{\cdot}0336t}$$

as a base-line for the further search for a relationship equation which must include T as well as t. At this point we must give some consideration to the appropriate functional representation of T.

67. If we were considering temperatures all the year round we might be inclined to depict the relationship between T and m' by a U-shaped curve; abnormally high temperatures, as well as abnormally low ones, may be expected to react unfavourably upon the rate of mortality. In prolonged heat-waves, associated perhaps with conditions of drought, bacteria breed, dust blows, flies multiply, milk goes bad, bowel complaints develop in the young and the infirm and old are liable to heat-stroke. But our attention is confined to temperatures in the March quarter and, variable as the weather of that quarter may be, the thermometer is never likely to climb high enough to set the mortality rate moving up the right-hand side of the U. For the purposes

Table 4. (See paragraph 66)

δ	A	B
−·490	15·32	− ·41
−·260	16·55	− 1·38
−·160	18·21	− 2·89
−·137	22·92	− 6·22
−·104	20·52	− 5·11
−·065	19·35	− 6·34
−·020	70·97	−54·09
−·011	64·12	−50·86
·025	−30·41	47·37
·045	− 3·59	17·76
·049	− 8·30	25·31
·100	3·37	12·07
·126	2·62	14·94
·161	5·30	12·32
·166	5·59	12·04
·167	5·65	11·99
·169	5·69	11·97
·169	5·70	11·95
·180	6·02	11·70
·184	6·09	11·76
·191	6·36	11·41
·198	6·48	11·52
·220	6·87	11·09
·229	6·57	13·85
·238	7·08	11·51
·243	7·41	10·68
·266	7·08	14·34
·273	7·71	11·25
·295	8·27	10·10
·300	7·56	14·71
·342	8·54	11·28
·376	7·94	17·30
·443	8·50	19·04
·584	9·24	24·79
·816	13·00	6·32

of practical analysis, therefore, we need a curve which is fairly steep for abnormally low winter temperatures but much less steep in the region of more normal temperatures. To define the form of such a curve, it is convenient to borrow the economic concept of elasticity. If we postulate a constant elasticity for m' with respect to T, we write

$$\frac{\partial m'}{m'}\frac{\mathrm{T}}{\partial \mathrm{T}} = -n, \quad \text{i.e.} \quad m' = \mathrm{C}\mathrm{T}^{-n},$$

where C is independent of T but may, of course, denote a function of the secular variable. Clearly a hyperbolic curve of this sort will satisfy the condition as to steepness mentioned above, although we could emphasize this characteristic by assuming

$$\frac{\partial m'}{m'}\frac{\mathrm{T}}{\partial \mathrm{T}} = -\frac{n}{\mathrm{T}},$$

in which case $m' = \mathrm{C}e^{n/\mathrm{T}}$.

68. The next step is obviously to test this *a priori* reasoning by calculating 'expected' values for m' from the secular relationship formula arrived at in paragraph 66. The constants in this formula were obtained from data relating to a period over which the average value of T was 41·4 (paragraph 64), whereas the average value of T over the period 1921–38 was 42·0. We might, therefore, expect that the values of m' for this period, as calculated from the secular formula, would be somewhat too high; and in fact the sum of these values turns out to be 177·97 as against 176·06 for the sum of the actual values, the difference of rather more than 1% corresponding roughly with the difference between 41·4 and 42·0. This at once suggests that we might try a combination of the secular and temperature relationships in the form

$$m' = (\text{expected value on secular formula}) \times \frac{41{\cdot}4}{\text{T}}. \qquad (3)$$

The results of this calculation are given in Table 5.

Table 5. (See paragraph 68)

Calendar year (1)	t (2)	Expected m' by secular formula (3)	T (4)	41·4/T (5)	(3)×(5) (6)	Actual m' (7)	(7)−(6) (8)
1921	−9	11·20	44·5	·930	10·42	11·08	+·66
1922	−8	11·02	40·9	1·012	11·15	11·13	−·02
1923	−7	10·84	43·6	·950	10·30	10·08	−·22
1924	−6	10·67	40·3	1·027	10·96	10·28	−·68
1925	−5	10·50	42·3	·979	10·28	10·37	+·09
1926	−4	10·34	43·6	·950	9·82	9·85	+·03
1927	−3	10·19	42·7	·970	9·88	10·10	+·22
1928	−2	10·04	43·3	·956	9·60	9·70	+·10
1929	−1	9·89	38·3	1·081	10·69	10·84	+·15
1930	0	9·75	41·5	·998	9·73	9·40	−·33
1931	+1	9·62	40·1	1·032	9·93	9·82	−·11
1932	+2	9·49	41·9	·988	9·38	9·50	+·12
1933	+3	9·36	41·7	·993	9·29	9·38	+·09
1934	+4	9·24	40·9	1·012	9·35	9·19	−·16
1935	+5	9·12	43·1	·961	8·76	8·88	+·12
1936	+6	9·01	41·0	1·010	9·10	9·08	−·02
1937	+7	8·90	41·5	·998	8·88	8·94	+·06
1938	+8	8·79	44·7	·926	8·14	8·44	+·30
		177·97			175·66	176·06	+1·94 −1·54
						Σ (A–E)²	1·31

69. The method of approach which has produced the series in column (6) of Table 5 may be described as a blend of general reasoning and *ad hoc* algebra. *A priori* assumptions have been given free scope and the ordinary 'fitting' conventions have been discarded. We have known from the start that our specification of independent variables was almost ludicrously incomplete: it would therefore appear to be unnecessary—if not, indeed, unsound—to impose the condition that the sum of the 'expected' values of m' over the period under review should equal the sum of the actual values, or to seek those particular values of the constants in the relationship equation which minimize the sums of the squares of the residuals. Apart from these

considerations of principle, there is the hard fact that if we had begun by postulating a relationship equation of the form

$$m' = CT^{-n}(A + Be^{-\delta t})$$

we should have been unable to solve for A, B, C, n and δ the equations resulting from the application of least squares procedure. Nevertheless, the fact that our calculated $\Sigma m'$ differs from the sum of the actual rates by less than ·25% justifies us in claiming that we have successfully reproduced the general level of the standardized mortality over the period under review, while our Σ (A–E)^2 is not greatly in excess of the values which result from the use of the two formulae of paragraph 61 (see foot of Table 2). We are also entitled to claim that inasmuch as our relationship equation has been derived by general reasoning based upon prior knowledge and common sense, the constants which it contains are *ipso facto* more intelligible than those which emerge from relationship equations which are chosen primarily on grounds of expediency, i.e. because they are amenable to a conventional technique.

70. We have not yet, however, quite exhausted the resources of general reasoning. The relationship $m' = CT^{-n}$ implies that the whole of the mortality rate is subject to the influence of March quarter temperatures, but it would be more realistic to suppose that only a part of it is so subject. To give effect to this supposition we should make part of m' independent of T and intensify the relationship between T and the remainder, by writing, instead of $m' = C\dfrac{41{\cdot}4}{T}$,

$$m' = C\left\{K + (1 - K)\left(\frac{\lambda}{T}\right)^{n'}\right\},$$

where $n' > 1$ and λ is a constant (cf. 41·4) specifically associated with T. In a normal year about 30% of the deaths occur in the first quarter, but it would probably be going rather too far to assume that no part of the mortality in the other three-quarters is affected by the temperature conditions of the first three months of the year. Let us, therefore, as a trial shot, take K = ·6 and—quite arbitrarily—adopt the reciprocal of T^2 instead of the reciprocal of T, retaining 41·4 as the temperature norm; so that our composite relationship equation becomes

$$m' = \left\{{\cdot}6 + {\cdot}4\left(\frac{41{\cdot}4}{T}\right)^2\right\}(A + Be^{-\delta t}), \tag{4}$$

where A, B and δ have the values assigned to them in paragraph 66. Before considering the results given by this formula we may notice that the effect of imparting greater realism to the temperature component has been to throw the relation between m', t and T into a form similar to that which we might have been inclined to postulate without any detailed *a priori* reasoning, viz.

$$m' = a + bf(t) + c\phi(T) + d\psi(t, T),$$

which could not have been 'fitted' by the method of least squares unless we knew not only the forms of $f(t)$, ϕ (T) and ψ (t, T) but also the values of any constants included in them. The form taken by (4) when the two expressions in brackets are multiplied together is shown by the four strips in Diagram 2, in which the heavy line represents the actual values of m'.

71. The values of m' computed from formula (4) in the preceding paragraph are shown in Table 6 alongside those obtained from formula (3) in paragraph 68

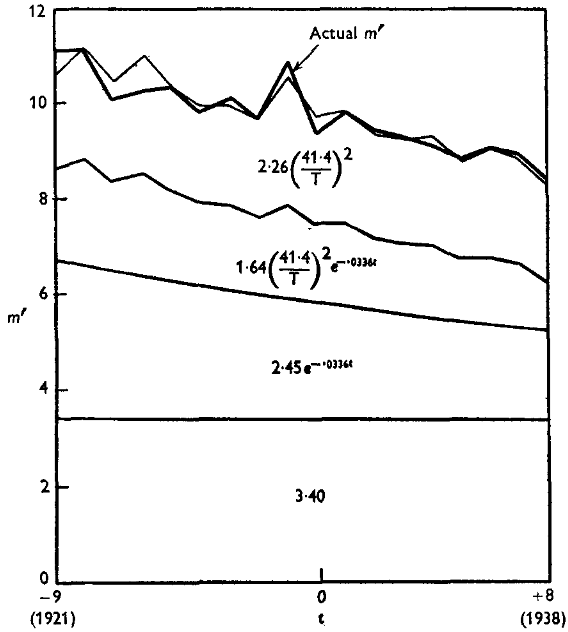

Diagram 2.

Table 6. (See paragraph 21)

Calendar year	Actual m'	Expected m' by formula (2)	(3)	(4)	A–E by formula (2)	(3)	(4)	T
1921	11·08	10·55	10·42	10·60	+·53	+·66	+·48	44·5
1922	11·13	10·96	11·15	11·13	+·17	−·02	—	40·9
1923	10·08	10·40	10·30	10·42	−·32	−·22	−·34	43·6
1924	10·28	10·77	10·96	10·90	−·49	−·68	−·62	40·3
1925	10·37	10·31	10·28	10·32	+·06	+·09	+·05	42·3
1926	9·85	9·99	9·82	9·94	−·14	+·03	−·09	43·6
1927	10·10	9·98	9·88	9·95	+·12	+·22	+·15	42·7
1928	9·70	9·76	9·60	9·70	−·06	+·10	—	43·3
1929	10·84	10·38	10·69	10·55	+·46	+·15	+·29	38·3
1930	9·40	9·75	9·73	9·73	−·35	−·33	−·33	41·5
1931	9·82	9·82	9·93	9·87	—	−·11	−·05	40·1
1932	9·50	9·44	9·38	9·40	+·06	+·12	+·10	41·9
1933	9·38	9·34	9·29	9·30	+·04	+·09	+·08	41·7
1934	9·19	9·32	9·35	9·33	−·13	−·16	−·14	40·9
1935	8·88	8·90	8·76	8·85	−·02	+·12	+·03	43·1
1936	9·08	9·06	9·10	9·08	+·02	−·02	—	41·0
1937	8·94	8·87	8·88	8·88	+·07	+·06	+·06	41·5
1938	8·44	8·36	8·14	8·29	+·08	+·30	+·15	44·7
	176·06	175·96	175·66	176·24	+1·61 −1·51	+1·94 −1·54	+1·39 −1·57	
Σ(A−E)²					1·06	1·31	1·02	

The table also recapitulates the results already shown in Table 2 for formula (2) (paragraph 61). It seems that the improvement of the temperature ingredient of the relationship equation was worth while; compared with formula (3) the results from formula (4) are closer to the actual values in thirteen cases and worse only in three cases. Formula (4) also gets closer to the sum of all the values and—for what it is worth in the absence of an exact equality between the actual and the expected $\Sigma m'$—the sum of the squares of the residuals by formula (4) is not only less than that by formula (3) but is also less than by the 'fitted' formula (2). This applies also to the sum of the deviations regardless of sign.

72. The effect of a change in the pattern of the temperature ingredient of the relationship formula is naturally greatest in the years when winter temperatures were abnormally low (1929) or abnormally high (1921, 1938). In applying the equation $m' = C\left\{K + (1-K)\left(\frac{\lambda}{T}\right)^{n'}\right\}$ of paragraph 70, values for the constants K and n' were chosen quite arbitrarily, and we were content to assume that the most appropriate value for λ was the 1901–40 average of March quarter temperatures. To investigate the matter systematically on the lines adopted for obtaining the secular constants A, B and δ we should need to form for each of the 18 years the equation

$$K + (1-K)\left(\frac{\lambda}{T}\right)^{n'} = \text{ratio of actual death-rate to expected death-rate by the secular formula,}$$

combine these in $\binom{18}{3}$ sets, devise a method of approximate solution for each set by making use of the assumption that, for any T within the range of our experiment, λ/T will not differ greatly from unity, and search for a modal cluster among the 816 sets of values of K, λ and n'. If we chose to retain the assumption that $\lambda = 41{\cdot}4$ the process would be less laborious to the extent that we should be concerned with two unknowns instead of three and 153 sets of equations instead of 816; moreover, the equations themselves would be easier to handle. But before embarking on this systematic inquiry we should be entitled to take note of the fact that in 1921—possibly on account of the abnormally hot dry summer—there was a severe epidemic of diarrhoea and enteritis which, on a rough calculation, increased the standardized death-rate for the year by something like ·30. When allowance is made for this the value A–E for 1921 in Table 6 no longer appears to be abnormally large.

73. It seems doubtful whether any refinement of the temperature ingredient on the lines suggested in the previous paragraph will greatly affect the substantial deviations shown in Table 6 for the years 1923, 1924 and 1930. On the very limited specification of variables adopted for the purpose of the analysis, mortality in these years (particularly 1924) was abnormally low; and we must be content to leave it at that. It would indeed be strange if over a series of nearly twenty years an analysis which took into account only three of the many factors which may conceivably affect the standardized death-rate did not break down at some point or other. Including the three unsatisfactory values, the mean difference, regardless of sign, between the actual rates and the calculated rates is less than 2 %; and we are, I think, entitled to derive some satisfaction from the fact that this result has been achieved by the diligent

implementation of common-sense considerations rather than by the blind application of a technique which has as its primary objective the satisfaction of 'fitting' criteria.

74. The temptation is strong to develop this very sketchy note on the trend of mortality over a selected period into an inquiry more elaborate and more worthy of the importance of the subject; but to yield to the temptation on the present occasion would defeat the object of this Appendix which, as stated earlier, is merely to illustrate some of the points discussed in the Paper in a short and simple fashion. For this reason attention has been confined to those factors which have the most obvious claims to be used as variables in an analysis of the mortality experienced within the period, with no thought of producing the most efficient formula for extrapolation purposes. For such purposes I would certainly prefer formulae (3) and (4) to formulae (1) and (2), on grounds which have already been stated; but it will also be appreciated from certain earlier remarks that if the object had been to produce a formula appropriate to the experience of an earlier period, the irregular incidence of diseases other than influenza would have had to be considered. Similarly, it is not to be expected that the mortality of the present decade could be calculated with any great accuracy from formulae (3) and (4) because, in the nature of the case, the provision made for secular improvement in these formulae can take no account of the effect of the recent introduction of penicillin and the sulphonamides into general medical practice. A further point is that since the eighties and nineties of last century the average temperature of the winter months seems to have shown a slight tendency to increase; in a more refined study designed to cover a longer period than 1921–38 some allowance might have to be made for this—possibly by substituting a moving average for the constant value of 41·4 which has been used in these calculations. Other considerations of a similar kind will no doubt occur to the reader; it is because I am anxious that he should not regard this Appendix as something which it does not pretend to be that I devote the last paragraph, as well as the first, to a statement of its restricted purpose and a recognition of its manifold limitations.

[The statements in paragraph 2 call for some qualifications in the interests of historical accuracy though they do not affect, in any way, the main thesis of the author. Most *early* mortality investigations were made solely or mainly with a view to the advancement of learning, not for the practical purposes suggested by the author. For example, the statistics on which the Northampton and the Carlisle Tables were based were collected and published ten and eighteen years, respectively, before they were used for life assurance purposes. A similar story was true of the early experiences based on assured lives. Finlaison's investigations into the mortality of tontines and of Government annuitants were primarily directed to proving that annuitants lived longer than was assumed and only later was his material used to provide a basis for the sale of annuities. From the earliest days mortality was studied in relation to many different variables. The remarks in the discussion on the relation of temperature to mortality could, for example, be closely paralleled by the article on the comparative mortality of different seasons of the year in the *Treatise* by Joshua Milne, published in 1815, which was a standard actuarial text-book for more than half a century. Eds. *J.I.A.*].

ABSTRACT OF THE DISCUSSION

The President (Sir George H. Maddex, K.B.E.) announced that the author, Mr Starke, was suffering from influenza and was not well enough to come that evening; he hoped the absence of the author would not diminish the liveliness of the discussion of what was a very debatable paper.

Mr K. Williams, in opening the discussion, said that nothing was more bewildering than the multiplicity of modern statistical methods. Workers in many different fields had, during the present century, elaborated devices which were specially adapted to the nature of their own particular material and which were intended to assist in the interpretation of the characteristics of that material. The body of techniques available was so large that few workers could do more than absorb and apply those which were already established in their own particular field. As always occurred after a period of rapid advance along different and specialized lines, there was need for consolidation of the progress made. Too often that task was nobody's business; the author's contribution was, therefore, eminently welcome.

The most fortunate workers in statistics were those in certain fields of scientific research. They had not only the most effective control of their material and of the design of their experiments, but also the further advantage of a fully developed and exact mathematical basis. That technical endowment made possible the accurate testing of the significance of their results (even though derived from very few observations) and also the estimation of any parameters inherent in their hypotheses. At the other extreme were those fields of economic and social research in which the data suffered from such deficiencies that the statistical material was described by the author as 'an omelette concocted from an unknown recipe'.

Between those territories lay, perhaps, that of the actuary. His land was ploughed very early in statistical history and it could not be said that the twentieth century had added much to the tools of those who tilled it. The author had explained the reasons for the apparent stagnation. The number of observations was so large that small-sample theory was quite unnecessary; the nature of the conclusions to be drawn from the data was generally known in advance and only the numerical values awaited determination. Those were found by methods which, though arbitrary in theory, were sufficient for the foreseen purposes of the investigation. In fact, there was still no accepted test of the graduation of a mortality table, nor an accepted system for the fitting of constants in, for example, a Makeham curve. Again, most actuaries applying a mortality table in practice probably never concerned themselves with the question of theoretical tests of agreement of observation with expectation. They nearly always had some mental reservation about the so-called 'expected value', for example that it included a little of that famous panacea, the 'safety margin'. The agreement they did look for was thus not with the true expectation but with an anticipated deviation therefrom of rather indefinite amount. Another example of the actuary's detachment from theory was revealed by Coward's recent pioneer investigation into the distribution of sickness, namely that the valuation of Friendly Societies had always been carried on without the benefit of any proper method of testing whether deviations from expected sickness lay within the limits attributable to chance, upon the hypothesis underlying the rates used in the valuation.

He believed, therefore, that there was scope in the mature actuarial field for a little experimentation with the new but well-tried methods of the biological statistician. They must not forget the fiasco of so-called 'spurious selection' in the $O^{[M]}$ tables, arising from the over-confident assumption of that practical homogeneity which was so well defined in the quotation from Hardy in paragraph 9. That particular trouble might have been avoided by a preliminary analysis of the results with reference both to duration and to year of observation—in fact, by a form of multiple regression analysis.

The author's concern, however, was with the other extreme of statistical territory. He wanted to see how the actuary, with his perhaps rather Victorian equipment, could help those economic statisticians whose fields were stony and whose tools were sometimes neolithic. Unfortunately, the actuary found himself forestalled by exponents of other and more modern panaceas, in particular of multiple regression analysis. That was a useful tool in its right place, particularly in the scientific field, and there it might be used successfully by a skilled computor without the need either for individual judgment in the processes or for the exercise of the actuary's inseparable companion, eternal vigilance. The uninformed use of that method, however, especially for forecasting, could easily lead to disaster. The actuary was, therefore, right to insist on getting the proper underlying form of relationship before deciding upon the fitting-technique. Where the form made it possible to use a mechanical method whose results did not depend on the computer, that was obviously economical of the expert's time; otherwise individual judgment must, of course, be retained. It was in the application of that criterion that he felt obliged to dissent from the views expressed by the author. He would relate his comments first to the paper proper and then to the example given in the Appendix.

The most general form for the multiple regression equation was that given in the footnote to page 194. The inclusion of combined functions such as $f_1(x_1, x_2)$ was quite permissible and overrode the limitation to functions of single variables mentioned in an earlier paragraph. The extension to curvilinear regression was obtained by writing $x_2 = x_1^2$, $x_3 = x_1^3, \ldots$, etc., and the processes remained identical. For curvilinear regression derived from a single observation at each of a set of equally spaced values of one independent variate, there was a great advantage in the powerful technique of orthogonal polynomials, By the use of that technique, it was not necessary to recalculate any of the constants on proceeding to a higher degree and it was easy to see at what point the process ceased to give any appreciable improvement to the fit and should, therefore, be terminated. That fact qualified somewhat the author's statement in paragraph 27 that the whole analytical technique could not provide an indicator of whether the relation was parabolic, though it remained true that the mathematical pattern should be determined in the light of all available knowledge. He believed, too, that the test of linearity of regression to which the author referred was one that required multiple observations of the dependent variate at each value of the independent variate and would not, therefore, be of use in the cases under consideration.

In regression analysis of the type being considered, the method of moments was identical with that of least squares, as might be seen by comparing the sets of equations which resulted when those two methods were applied to curvilinear regression with one independent variate. The author's doubt about the use of moments for the multivariate case appeared to be simply a confusion in terminology. It was pointed out in the paper that the results of the least squares procedure were also equivalent, in the case considered, to the results obtained by calculation of the multiple correlations in the ordinary way. On the reasonable assumption that the errors were normally distributed with a variance constant for all values of the independent variate, the results coincided with those of yet another method of estimation, that of 'maximum likelihood' due to R. A. Fisher, the omission of whose name made the author's paper unique amongst recent papers on statistical subjects. In the view of many, the equivalence with the maximum likelihood procedure probably represented the best justification of the least squares procedure; otherwise it seemed to be entirely empirical. The meeting that evening was not the place to discuss the abundant merits of the maximum likelihood method, which were leading to its increasing use; they would be found well set out in Kendall's *Advanced Theory of Statistics*, Vol. II, and recently in the literature of their own profession in *The analysis of heterogeneous mortality data*, *J.I.A.* Vol. LXXIV, pp. 94–112, a paper by Leon Solomon which had excited interest among scientists generally as a valuable non-rigorous introduction to the subject.

The practical justification of the least squares procedure, even in the exploratory stages, was that it avoided personal bias and such extraordinary and laborious artifices as were contemplated by the author in paragraphs 39–42, in which the fitting of three constants to a series with, say, fifty terms (a number quite common in actuarial practice) would present him with the possibility of 19,600 sets of equations to be solved or sampled! Even if the least squares equations were formally insoluble, the process of trial and error need not be blind if it were remembered that the partial differentiation of the normal equations with respect to each of the parameters in turn gave a direct measure of the expected effect on each equation of small changes in the parameters. Thence, a set of simultaneous equations for the corrections required to the trial values of the parameters would be obtained. That was the clue to an iteration process much more easily contemplated than the precipitous incline of a table of the binomial co-efficients $\binom{n}{r}$. The process made practicable the fitting of Makeham curves by least squares methods. A first application to the data used in paragraph 66 of the Appendix reduced the sum of deviations from the high value of 0·32 to the much smaller value of −0·04, and at the same time reduced the sum of squares of deviations from 0·48 to 0·44.

Another case in which difficulty could be removed was the solution of the equation in powers of c given in paragraph 31. With a knowledge of the approximate value of c, a calculating machine and a good logarithm table, the result could be obtained in a few minutes by trial and error. The limit to the accuracy of the solution was simply that of the aids to calculation employed, which should certainly go at least as far as the practical problem required. The search for explicit expressions for the roots of equations of high degree was surely academic since aids to calculation were readily available to those with a practical interest in the matter.

The author's special knowledge soon reduced the problem set out in the Appendix to that of relating the death-rate in the years 1921–38 from causes other than influenza to the variates, time t and first-quarter temperature T. He first set up two possible forms of linear regression equation based on the data itself and found the required parameters by least squares procedure. The author then criticized the implications of those hypotheses for purposes of forecasting and proceeded by gradual empirical steps to arrive at a fit which was as good as those obtained from the regression equations and related to a more readily comprehensible hypothesis. In the process he took in extraneous data from a further range of years as long as that primarily considered; he gave to the temperature-effect a form which involved a degree of elaboration that the data would not support and that also misrepresented the magnitude of that effect; and he provided an example of the pitfalls of what might be called the $\binom{n}{r}$ procedure. All the difficulties might, in the opinion of the speaker, have been avoided by adjusting the form of the regression equations to overcome the drawbacks which the author had observed to be inherent in the two forms of relation which he had used [equations (1) and (2) on p. 208].

Table 4, like the Great Pyramid, was most impressive; but he was glad not to have toiled in the construction of either! The four sets of values between the lines were all derived from the four sets of three observations obtainable from the entries appearing in Table 3 against the years denoted by $t=1, 2, 7, 8$. The other three years had not been allowed to give any weight at all to the analysis. Any three values would, of course, be amenable to the form of relation postulated. Further values might or might not agree well with the curve determined by the three first chosen. It so happened that the author's four values shared the fortuitous property that any one of them would fit almost exactly on the curve determined by the other three. He could not see that those values had any other merit, and it was unfortunate that they happened also to represent the four years of lowest temperature, excepting only $t=4$.

He feared that the author's discussion of the effect of temperature was, to quote him in another context, 'rather quaint'. The range of variation of T from its mean value was less than 10% of that value, and the values of c in both equations (1) and (2) showed that the change in m' for increase of one degree in temperature was about −·13 to −·15. The maximum effect on the predicted value—in other words, the effect due to a change in temperature of about four degrees—was, therefore, ± 0·6, about 6% of the result. No significantly wider variations in temperature were to be expected in future. Now, over such small proportionate variations in T and m' there was nothing to choose between the representation of temperature by a linear correction, by the forms of paragraph 17 of the Appendix, or by the further complications postulated in later paragraphs; all such methods must give, over the relevant part of the range, results practically equivalent to linear variation. In other words, the change due to a variation of, say, four degrees in the temperature would be indistinguishable from twice that due to a change of two degrees. Any one of these methods should, however, be properly fitted if it were not to distort the magnitude of the temperature effect. The factor 41·4/T gave a decrease of about 0·25 in the standardized death-rate for each degree of increase in temperature; this was almost twice the variation revealed by the two regression analyses which were independent of personal bias and predilection. The order of this mis-statement of magnitude was also confirmed by substituting in paragraph 18 the true mean of the four values used from Table 3 (namely 41·15) for the overall mean of that table (42·0). It would then be found that although the difference in the sum of m' was rather more than 1 % that in T was just 2 %, instead of approximately the same proportion. The final formula for the temperature effect given in equation (4) also over-adjusted for temperature, for the decrease which it gave in the standardized death-rate per degree of temperature in the range of expected temperature variation was about 0·20. The only reason why the effect of that crudity of treatment was not particularly apparent in the results set out in Table 6 was that the temperature effect was of little account beside what might be called the natural awkwardness of the particular years for which all the formulae produced bad results. In those years the variations could obviously not be accounted for by the particular independent variates considered. Whatever view might be taken about the effect of an incredibly high or incredibly low first-quarter temperature, account need not be taken of such variations simply because they were incredible. Considering only such variations in mean temperature as had been experienced in the past, the effect of temperature upon the death-rate would be recognized as sensibly linear over the range of variation observed or expected. The proper criteria were therefore satisfied by so representing the effect in the regression analysis. Were there, actually, strong reason to believe that the effect of a 10 % variation from the mean temperature was very much greater than twice that of a 5 % variation, then the representation of that effect would require formulae involving a very much higher value of n or n' than was used in equations (3) and (4).

It would be found, with regard to the time-effect, that both the regression methods agreed in producing an average annual change in the standardized death-rate of about −·13. Though the method by which the Makeham constants were derived was based on other years than those of the primary observations, it, too, gave a similar average change over the eighteen years. It was, indeed, very reasonable to introduce a limit to the possible fall in mortality, but that could be achieved by other methods than the introduction of the Makeham form with negative exponent. In fact, if temperature effect were ignored and all variates were related to their means, as in equation (1) in paragraph 26, it could be postulated that the logarithm of the standardized death-rate decreased from year to year in geometric progression and m' would decrease to a limit as in the Makeham form. The approximate annual decrease in m' over the period of the observations was already known as, also, what was a reasonable relationship between the initial and the limiting values of m'. From this knowledge it could be seen at once that the geometric progression would have a constant proportionate decrease of about $2\frac{1}{2}$ % per annum. Using compound interest tables to bring in the temperature-effect in its simplest form, equation (2) in paragraph 61 could be replaced by the following equation:

$$\log m' = \log a + b v^{t}_{2\frac{1}{2}\,\%} + cT.$$

The values obtained by fitting the equation by the usual least squares technique compared well with any set in Table 6. The positive deviations in m' totalled 1·52; the negative 1·42; the squared deviations 1·015. The method seemed to have the double advantage of representing a reasonable hypothesis as to future mortality in a form amenable to regression technique. At a standard temperature of, say, 42 degrees, the values of m' obtained during the next hundred years were in close accord with those of the author's Makeham formula and the limit to which the standardized death-rate might fall was 5·23 (compared with the author's 5·66). If it were desired to postulate a different order of relationship between the initial and ultimate death-rates, the rate of interest used could be varied accordingly. If, however, it was not considered legitimate to make any such restraining postulate, then all the parameters (including the rate of decrease) should be fitted by the iteration process to which he had referred.

That concluded his remarks on a most interesting paper. He would like to express his thanks to the author for rousing him out of the armchair into the kitchen, away from the Third Programme and the passive perusal of *The Times* to the breaking of eggs and the active tasting of omelettes.

Mr W. Perks said that after the brilliant opening speech there seemed very little left—for some of them at least—to say.

Any actuary who might, at that time, be endeavouring to write a paper on a statistical subject could hardly avoid exhibiting the limitations of his knowledge of mathematical statistics. The author had not escaped the danger, but there was a wealth of wisdom in his paper. He had not allowed mathematical analysis to blind him to the plain facts of arithmetic, and somehow he had usually managed to get a pretty reasonable answer to the problems he had posed.

Particularly pleasing were three paragraphs towards the end of the paper—paragraphs 46, 47 and 48. In paragraphs 46 and 47, the author said in a much better way what he himself had said at the previous meeting. In paragraph 48 the author made an appeal to actuaries to endeavour to extend the techniques they used, an appeal which the speaker supported.

He wished to amplify a little what the opener had said about the Appendix. The author had used the mean temperature in the Fahrenheit scale as the measure of the temperature effect. The particular thermometer used would not matter if the temperature effect were represented either in a linear, a parabolic or an exponential way. But when the variable $1/T$ was used the particular scale of measurement was critical, and the whole of that section of the paper would have been different had the author used Centigrade instead of Fahrenheit.

The opener had pointed out that owing to the range of values of T round the mean value being small the device of using $1/T$ instead of T did not really depart much from linear variation. This of course, depended on the fact that the Fahrenheit zero was 32° below freezing-point. Actually, by expressing T as (41·4 + the deviation), and the deviation by $(41{\cdot}4\delta_T)$ the fraction $41{\cdot}4/T$ became $(1+\delta_T)^{-1}$, where δ_T was always small. The expansion of $(1+\delta_T)^{-1}$ showed that $41{\cdot}4/T$ did not depart significantly from linearity. In formula (4) the author used $(41{\cdot}4/T)^2$. By expanding $(1+\delta_T)^{-2}$ it could be seen that there was again no significant departure from linearity.

That being so, he had thought it worth while to take formula (3) and to substitute $A+BT$ for $41{\cdot}4/T$. The effect of fitting by moments for $A+BT$ was to produce a 'graduation' which, like the opener's alternative 'graduation', compared favourably with any of the 'graduations' in Table 6. All these 'graduations' brought out the fact that the mean temperature explained only a fraction of the fluctuations in the standardized death-rate, probably less than half. The random variation in the standardized death-rate was extremely small because the rate was based on the deaths in England and Wales in a given year, a very large number. The author's result thus left unexplained about half or more of the fluctuations. Experience of British mortality from year to year showed that there was more in the effect of the weather than that. He could only assume that the mean temperature in the first three months of the year was not the best measure of the temperature- or weather-effect.

It might be that peculiar things happened as between temperature and mortality—so peculiar that they were not amenable to mathematical representation; but his personal experience was that if the temperature was 60° for two days running and then dropped to 25°, then shot up again to 55°, then fell to 35°, those were the conditions in which he caught cold after cold—and he felt himself lucky to escape pneumonia.

Such conditions might lead to a normal mean temperature, and it seemed that what was wanted to measure the temperature-effect on mortality was some indication of the variation of the temperature in the first three months. No doubt the mean temperature was a factor, and it occurred to him that the coefficient of variation of the daily temperatures might be a single figure that combined both effects—the variation of the temperature and the mean temperature in those three months. He would also not exclude the possibility of auto-regression, such as might happen if a bad winter weeded out the weak lives and left a relatively select group for the following winter.

Mr B. Benjamin, notwithstanding the obvious sincerity of motive of the author, was inclined to regard the paper, like laughter in the dentist's waiting-room, as being rather forced. It was forced in the attempt to find an excuse for the distinction between the conventionally 'actuarial' past of the profession and its expected 'statistical' future. (He was using 'actuarial' and 'statistical' in the same sense as the author.) And it was forced in that the hypotheses used by the author were unnecessary to explain what had happened. Until recently actuaries had not been found playing with the tools which modern statisticians found so useful. But how long had the statisticians themselves been so recondite in analysis? The average statistician of twenty years ago had used very simple tools. The statistical science was rapidly advancing, and if actuaries were caught up in the advance it was not something for which they should shyly apologize but something to which they should energetically contribute, as, indeed, actuaries had contributed in the past.

In paragraph 3 the author said:

> 'a good deal of statistical research is undertaken in the knowledge that, if it produces positive results, there are practical uses to which those results can at once be applied'.

That was surely an under-statement. Most researches set out to answer a difficult question which must be answered to resolve some impasse in applied science. The author had emphasized the connexion between the conventional actuarial technique and the unique accuracy of assurance mortality data to which it was applied. But there was the reverse side of the penny. Why had actuaries neglected the study of those aspects of their material which were not so blessed with precision, e.g. causes of death or socio-economic selection? He did not think that the avoidance of those variables by straining after homogeneity was so much a conscious as an unconscious recognition that the available tools were then inadequate. He did not think that actuaries of yesterday were less beset by problems of multivariate analysis than the economists before what might be called the *Stone* age, who also failed to resolve those problems.

In the fantasy in paragraph 18 there was envisaged the nightmare of adequate economic data but no vital statistics. Yet it was not just an accident that life had been the other way about. People had always been afraid of death; but, in a *laissez-faire* economy, they had not worried quite so much about the price of tobacco. The economist was catching up with the actuary in the precision of his data. To borrow the author's analogy, the omelette was not made of shell eggs, however mixed, but of dried eggs. Life was hard and they were faced with the stark necessity of abandoning *laissez-faire*. The economist of tomorrow might have such excellent data that the author would insist on his being content with an early actuarial textbook.

In paragraph 36 the author might have emphasized more explicitly the criterion of residual variance in deciding whether further variables should be sought.

Mr Perks had already referred to paragraph 46, and he would like to add his own praise to the author for his distinction between interpolation and extrapolation and his warning of the dangers of the latter operation. Prof. Greenwood had illustrated this very well in a recent discussion at the Royal Statistical Society: he pointed out that if infant mortality were expressed as a function of several variables of which one was fertility, it was clearly useless to extrapolate by putting fertility equal to zero, because infant mortality would also be zero. Yet something as bad as that had been known to happen.

If parts of the paper had made his hackles rise, most of it had delighted him and he had found the main exercise in correlation interesting and instructive. Sulphonamides had been introduced in the middle 'thirties, too late to affect the pre-1938 picture. Since then, as the author remarked, sulphonamides and antibiotics, like the Ministry of Fuel, might have reduced the power of temperature. That brought him to the warning that an analysis of the kind the author had made should not be attempted without medical guidance in order that the biological implications might be as fully understood as, for example, in market research the economic implications were fully considered. Otherwise the choice of variables might become entirely unrealistic.

Mr J. G. Day welcomed the idea of simple, common-sense methods. To the ordinary person, and even to the ordinary actuary, the standard method and formal solution, even of ordinary problems, were very difficult and very laborious. The author must have had some difficulty in solving thirty-five sets of equations; it would have taken him some time! But those simple methods were used in practice, and even in a scientific laboratory, and it was time they appeared in a learned journal. Also in comparison with the multiple regression analysis the ordinary association table with χ^2 test was fairly easy and very useful.

The actuary used the variates age and duration, he suggested, because they were the easiest to use and the only ones not affected by other variables. They remained definite whatever the office did.

The author had perhaps implicitly assumed that the actuary was a statistician. But an actuary was not quite a statistician. He had always to remember financial considerations. He did not set out to obtain the true and ideal answer. The statistician applied purely methematical tests to know whether the answer was significant. The actuary, on the other hand, used his judgment and wanted to know his errors and, if possible, to have them on the right side as a margin. He could never divorce his answer from his financial considerations.

Mr C. D. Sharp dealt with a point which though of minor importance in the paper was of interest to many actuaries, namely, the assumption that for life assurance purposes mortality statistics should be analysed by age and by duration. It was his belief—a belief supported by a certain amount of evidence—that in many parts of the world the social standing of the assured was of more importance than duration and possibly than age. The variation in a table of twenty-year endowment assurance premiums was comparatively small over a wide range of ages, whereas the variation in the premiums when proper allowance was made for social status could be much more.

Broadly speaking, social status might be correlated with the sum assured, and he suggested that the difference between the premiums brought out by the A1924–29 Light Table and the A1924–29 Heavy Table could be taken to indicate the possible range in Great Britain. An analysis by sum assured of the mortality experience of assured lives might well provide evidence of considerable interest.

The idea had been put forward that by selection mortality results could be produced for other races similar to those experienced in the United Kingdom. Though such a course might be practicable in countries with a similar type of society to Great Britain, it was certainly not true of the Asiatic countries. The point he had been trying to bring out was that whereas mortality had conventionally been analysed by age and duration, factors such as social status might have become more important because of changes in the rates of mortality. That applied particularly to those countries where there was a wide variation between the mortality of the upper classes and the mortality of the general population. Apart from general considerations he had seen statistical evidence supporting that point of view and he considered it was an aspect of mortality investigation which required further consideration.

Mr H. W. Haycocks did not consider that the author had laid sufficient stress on objective or purpose in his attempt to explain the differences between the statistical techniques used by actuaries and those used in other fields. The author had emphasized the actuary's freedom to select and adapt data, and in dealing with that he had said—and the statement was not easy to understand—'expressed mathematically, . . . the actuary can go a very long way in the direction of isolating his variables'. All that that seemed to mean was that the actuary could measure such variables as age and duration accurately and could then express the rate of mortality as an approximate function of those variables. But in most sciences that was possible for some variables. Often, however, they were not the important ones and frequently it was impossible to obtain what might be termed 'experimental isolation' to distinguish that operation from the author's 'mathematical isolation'. In scientific method it was 'experimental isolation' that was important. It was clear from the immense literature on spurious selection that 'mathematical isolation' was not enough and could be very misleading. In the sociological sciences it was very difficult, often impossible, to obtain an exhaustive set of independent variables. Even when that stage was reached the statistician required that the scientist should provide him with alternative hypotheses which could be tested. The function of statistical analysis in, for example, econometric research was that of an intermediary between a general theoretical hypothesis and the directly observable facts. The progress of such research would depend as much on successful and non-statistical research for suitable analytical tools as on a final statistical manipulation of the observations.

It was essential to notice the difference between factor analysis and multiple regression analysis. In the latter an equation was obtained showing the relation between a specified variable and a specified set of other variables. In so far as the variables were independent then the regression coefficients indicated the relative weights of the explanatory variables. As the opener had clearly shown, the analysis in the Appendix did not do that. There was one point in the opener's statement which he had been unable to appreciate. If the regression equation contained an explanatory variable which was a function of two or more variables, what interpretation could be placed on the regression coefficients? The effect of a single variable would be contained in several coefficients.

Factor analysis was more complex. It had been applied in much detail in psychology. The observed data were a set of scores obtained by a group of individuals in a battery of tests. The analysis attempted to reduce that structure to one based on only a few factors—say two or three—and it was hoped that those factors would be more fundamental in that they could be shown to correspond with some physiological phenomena. If that was not so, then factor analysis simply gave a more economical and convenient classification of the data. In those cases which he had seen where the method had been applied to mortality, only a more economical classification had been obtained, and even that was dubious. The data were a set of standardized mortality rates and a corresponding set of index numbers relating to unemployment, social class, housing density, latitude, etc. Factor analysis then showed what was obvious, namely that several of those factors were highly correlated and could be replaced by a single index of social conditions. It seemed to him doubtful, however, whether such an omnibus and vague index was helpful to those who were seeking to control mortality rates.

A great deal of scientific activity was concerned with finding controls over the environment and over themselves. For some reason it might be desirable that a certain phenomenon should be controlled, but very often it was impossible to do this directly. For example, a disease caused by the sale of a dangerous drug could be controlled directly by prohibiting the manufacture and sale of the drug. Generally, however, such a simple procedure was not possible. In the case of, say, poliomyelitis or cancer it was necessary to analyse all the circumstances of the disease in order to ascertain whether among those circumstances there were any factors highly correlated with the intensity of the disease and at the same time subject to direct control by a human agency. Much scientific research was of that nature and often it was necessary to use statistical techniques.

In Life Assurance sufficient funds had to be accumulated to pay claims and to cover costs. That required a stability in the mortality of the assured population, which could be obtained by a fairly simple control over entrants. Having by experience found efficient controls for the purpose the actuary based his premiums on a reasonable assumption about mortality. The scale of premiums had to be safe, competitive and equitable to policyholders; further, the form ought not to tend to weaken the control of entrants by introducing a selective effect. There was nothing very fundamental about a structure designed for such a limited purpose. Unlike the medical man or social worker the actuary did not seek a fundamental set of factors to explain mortality. Its advantages with regard to statistical techniques arose from the facts that his objective was limited and that he could control his population. He had not the invidious job of making accurate forecasts, otherwise he would have had to look for other techniques. In the case of annuity business where other techniques had been used the actuary found himself in the unfortunate position of other workers who had to make forecasts.

Techniques depended very much on purpose. Refinement was a waste of effort if it was unnecessary. However, he was not a utilitarian. Like the author he would pursue science for its own sake and would encourage other actuaries to do likewise. But then they would have to use new tools for they would be tilling new ground.

Mr H. A. R. Barnett said that he wished to make one or two comments, largely on paragraph 30, though they might not seem very relevant to the purposes of the paper.

In the first place, the author seemed to assume that the only way to fit a Makeham curve to mortality data was to find a trial value of c and then by the method of moments or of least squares to find the corresponding values of A and B. It was not really necessary to do that to find A and B.

Suppose, for example, crude values of μ_x at quinquennial intervals were available. Having arrived at a trial value of c, it was then possible, by means of the expression

$$B = \frac{\mu_{x+5} - \mu_x}{c^x(c^5 - 1)}$$

to arrive at a succession of crude values of B. If those fell fairly close to each other, a trial value of B could be derived, employing, if necessary, a suitable series of weights. Having, then, trial values of B and c it was possible to work backwards and to arrive at what might be called crude values of A ($\mu_x - Bc^x$). That sounded very laborious, but in fact it was not. Once c had been derived, B and A followed very quickly, and the three could then be put together and the usual tests applied to the formula. By looking at the results one or all of the constants could be improved by a hand-polishing method, which was similar to the method used in a graphic graduation apart from the fact that the graduated values followed Makeham's law. He agreed that the method was open to the objection that it was difficult to see whether the best formula had been achieved. He suggested that the method could be carried a stage further by applying a Makeham minimum χ^2 method; not the method mentioned by Dr Pollard, *J.I.A.* Vol. LXXV, p. 159, which merely produced the minimum possible χ^2 consistent with zero totals of deviations and accumulated deviations, but a method producing an absolute minimum χ^2.

He had been trying to develop the method and had found it was possible to arrive at an absolute minimum χ^2 by a trial and error method, provided the first trial was a fairly close fit. He suggested that the way to arrive at that first trial was to hand-polish by the method just described, which Mr Barley had privately christened the 'flexible method'. That minimum χ^2 method would, he thought, overcome some of the author's objections to the usual methods of fitting a Makeham formula.

Mr B. Robarts thought he was what one of the previous speakers had referred to as an ordinary actuary, in that some of the developments of modern statistical technique did not come altogether easily to him. Nevertheless, he felt that a paper of the kind being discussed that evening, which tried to relate those developments to the more accepted actuarial techniques, was of great value. It was of value because it caused them to think where they were going, what use they could make of modern statistical techniques, and perhaps more important still what contribution they could make towards them.

It would be gathered from what he had said that he was not going to venture into the technicalities of the paper, but he would like to make one or two comments on part of its general subject.

He was interested particularly in paragraph 4 where the author said:

> 'Let us consider first the position of the actuary in relation to his data. In the organization which collects these he is a responsible official whose ideas and requirements are treated with respect—probably, indeed, with something approaching reverence.'

A little later on the author referred to the fact that the actuary was normally dealing only with two variables. He wondered whether the reverence would be in direct proportion to the number of variables, but for all that he did think there was, underlying the closing words of the quotation, a grain of truth, which was worthy of investigation.

He had been fortunate, during the war of 1939–45, to have had some experience in carrying out trials on anti-aircraft equipment which involved statistical analyses. At the beginning of the war statistical analysis did not go beyond the stage of mean and mean deviation, but by the end of the war very considerable strides had been made in adopting many of the modern techniques, and the transition was not easy. On the one side, there were men who were able to carry out the practical work and produce the results but were not trained in modern statistical methods; on the other side, there were statisticians who were competent to analyse but quite unable to carry out the practical work. The difficulty was to find a common meeting ground for them. Any experiment could be divided into four stages. First of all, there was the design of the experiment; then the practical work; then the analysis of the results; and finally, their interpretation. The practical men could deal with the second, the practical work; the statisticians could deal very well with the third, the analysis of the results. But in the first stage, the design of the experiment, and in the last stage, the interpretation of the results, they had to meet together, and it was there that the difficulty lay.

Actuaries were perhaps a little more fortunate in that they tried to combine both persons in one, but at the same time had a harder task, because they had to keep their technical equipment up to a high pitch of efficiency and yet, so to speak, keep their feet on the ground and look after the interpretation of the results when they had obtained them.

He felt that they had perhaps not always been as careful as they might have been in the matter of interpretation, rather contenting themselves with the design of experiments, the practical work and the analysis. The question of interpretation should receive more attention, because it was on the interpretative side of their work that they would be judged. They should, in fact, after they had counted the trees, stand back and look at the wood and then describe what they saw in the simplest possible terms. Perhaps then they might receive slightly less reverence but—he was quite sure—added respect.

Mr F. M. Redington, in closing the discussion, said the author had, as it were, taken them by the hand for a leisurely stroll through his garden. It would be churlish to make the sort of comment that might be made at the Chelsea Flower Show. In that spirit the meeting might agree with many of the author's general sentiments, one or two of which he himself would elaborate.

The author, in his useful discussion of the relationship between the dependent and the independent variables had made the interesting suggestion in the footnote to paragraph 20 that perhaps there should be no constants in the formula and that all the elements should be variables. In that he was probably right. To be philosophical, in the author's own vein, all phenomena were to some extent interrelated. It was true that in any particular problem the vast majority of phenomena could be discarded as having too remote a relevance. But there would remain a considerable number of phenomena which were immediately relevant. The personal history of every individual in a mortality experience was relevant to the statistical outcome. It was, of course, out of the question to deal with every such element. What was assessed, in fact, was not which elements were relevant and which non-relevant, but which were random and which non-random. The quality looked for was 'non-randomness'. The elements into which a mortality experience was divided—duration, age, or, it might be, size of sum assured—were chosen because they were unlikely to enter into the experience in a random way.

In paragraph 23, the author had summarized Stone's antecedent considerations, one of which was 'to bring in as far as possible all the important influences on the dependent variable'. He would like to make a slight correction there. What should be, and indeed instinctively was, brought in, as far as possible, was all the important *non-random* influences. Indeed, though the words homogeneity and heterogeneity were often used in actuarial literature, the word 'homogeneity' could really be dropped. Heterogeneity could never be avoided. The aim was perfect heterogeneity; it was imperfect heterogeneity that required separate analysis.

The next general point, on which the author had spoken wisely, in particular in a very excellent passage in paragraph 47, was the introduction of time itself as an independent variable. The fact was that while a time record had many attractions, it was comparatively seldom that time itself was a direct, operative, functional variable—that the mere passage of time was itself causing the variations which were perceived in a time-record. There were, of course, exceptions. There were such things as the effect of astronomical and biological changes due to the passage of time. But the scale of those influences was too large to influence the problems with which they were generally concerned. There were other and more practical examples where time itself was an active agent. An obvious one with which perhaps they were all too familiar (because it might lull their senses to the dangers of using time as a variable) was the increase of mortality with age. It might be that even there it was not so much the passage of time itself that was important as the amount of wear and tear on the body, the number of colds one had suffered, and so on. But all those were so closely correlated with time that the study of mortality according to age had been both profitable and perhaps essential.

When time-changes in fertility or in generation mortality were studied with a view to making forecasts, the very different and quite unwarranted assumption was made that time was the operative factor in causing the observed changes. The very word 'forecast' implied that time, being itself the causative element, would continue to cause the same changes and that a formula which had fitted in the observed past would apply in the future.

Of course, it could not be emphasized too strongly that a forecast, such as that made for the $a(f)$ and $a(m)$ tables, did not assume time to be a causative agent. The process was one of extrapolation: certain social and economic conditions in the years 1880 and 1910 were reflected in the observations and it was assumed that corresponding changes in those social and economic conditions would continue.

It was a pity, therefore, in view of the author's excellent remarks in paragraph 47, that he should, in the Appendix, have chosen as his object of study an example in which 'the practical course is to treat the various possible specific factors as a conglomerate epitomized by time itself'. As a result, the whole of the Appendix had empiricism as its foundation. He had no complaint at all with that, but the sort of structure that the author placed upon that empirical foundation seemed a little ornate.

Most of the points he had intended to raise on the detail of the Appendix had already been made by other people. Moreover, the author had stressed that the Appendix was an illustration only and it was perhaps more fair to concentrate on methods rather than on detail. The outstanding feature of the author's methods was his use of the cluster method and he, the speaker, agreed with some of the criticisms that had been made. The opener's remark about 'the precipitous incline of a table of the binomial coefficients' was pertinent. It would be quite impossible to use the cluster method for an extensive experience, except perhaps by sampling, and where that would lead he would not like to say. But there were perhaps occasions on which the cluster method might be useful.

If, for example, the war years were included in an experience, any graduation which used an average value for the constants A, B and c calculated by the usual methods might be completely wrong; whereas the cluster method might give a reasonable answer. Of course, it was much simpler to leave out the war years, but there might be certain large influences whose presence might be suspected though they could not be specified.

Broadly speaking, where there were abnormalities in the experience of the 'black or white' type, for example 'war or no war', the cluster method could be useful. But if the abnormalities were of the 'grey' type with degrees of variation—a good example was temperature in the Appendix to the paper—he would be suspicious of the cluster method.

He would like to conclude on a note of complete agreement with the author. In probing into the unknown, which was their purpose in trying to bring some logical order into a body of crude data, all possible sources of knowledge should be taken account of and conclusions should be seasoned with a considerable dash of self-criticism, not to say scepticism.

A short quotation from Stone's paper with which he thought the author of the paper would sympathize was apposite:

> 'There has been a tendency, I think, especially among the users of the results of statistical economics, to underestimate the difficulties of induction. At the same time, practical questions have to be answered and we must hold fast to the most substantial straws we can find.'

The President, in proposing a vote of thanks to the author, said he was particularly sorry—as they must all be—that Mr Starke had been unable to attend the meeting and to reply in person to the discussion.

It was clear that over a large part of the paper the author was thinking aloud and looking for guidance and assistance in a country which was not entirely familiar to most of them. Frankly, if he himself were asked to give directions, he would have to say, 'My dear fellow, I am a stranger here myself'. But, evidently, a fair proportion of those in the Hall that evening were not quite such strangers.

He was sorry also that Mr Stone, who had hoped to come, was unable, on account of illness, to be present. He, more than anybody, was a practised expert in applying to economic statistics.the particular type of analysis to which the author had referred.

On the subject of objectives, it was clear, of course, that in his paper the author was just wanting to know, to understand. That was a scientific attitude, which had little to do with the actuary as a computer of premiums or as a manager of an insurance fund.

What appealed to him as one of the fundamental points was the emphasis of paragraph 47; the idea that before 'goodness of fit' could be discussed as a criterion we needed to be confident that the relationship equation was of the right pattern. That seemed to be the crux of the matter, and it involved a considerable act of faith. The statement, in paragraph 36, that

> 'our first rough approach to the relationship equation might be an expression involving only the major independent variables, the minor ones being regarded for the time being as hidden in the constants by which the variable terms in the equation are connected'

was not a question-begging statement and particular aspects of that had been mentioned in the course of the discussion.

The script would be passed to the author who would have an opportunity to reply in writing.

Mr Starke subsequently wrote as follows:

I have read the discussion with great interest, though I feel bound to say that I should have found it even more interesting had it been concerned rather more with the ideas in the paper and rather less with the calculations in the Appendix. It seems that I did not succeed in making it clear in the first and last paragraphs of the Appendix that the object of the calculations was not to find the particular equation connecting m', t and T which could be regarded as the most satisfactory from all points of view, but merely to illustrate by reference to an actual set of figures one or two of the problems which seemed to me to be inherent in the whole subject of multivariate analysis.

For this purpose it was necessary for me to find not less than three statistical series (preferably relating to a subject of some interest to actuaries) and to construct around them, by general reasoning, a theory leading to a relationship formula which was *not* amenable to the normal methods of analysis. Mr Kennedy Williams, on the other hand, set out to find, and succeeded in finding, a formula which *was* amenable to established technique as well as reasonable from the point of view of *a priori* considerations. Even if his formula had occurred to me, I could not have used it without defeating my object.

The example I chose was, I am afraid, not a very good one; it would have been better to invent some data for the purpose. Though I adhere to the view that, considered in the abstract, the relationship between T and m' (corrected for the secular trend) ought to be curvilinear, I admit that there is no very convincing evidence of non-linearity in the limited range of values provided by the data. I agree with Mr Perks that a temperature index which took rapid changes into account might well provide a more useful variable than the mean values of T over the whole quarter. In constructing such an index it would be necessary, I think, to have regard to the normal seasonal trend between the beginning of January and the end of March. I had some such idea in mind when I began to plan the Appendix, but since the only data conveniently to hand were the mean monthly values I persuaded myself that for the limited purpose I intended the Appendix to serve it was not really necessary to pursue the idea.

I think, too, that there may be a good deal in Mr Perks's suggestion that a bad winter may weed out some of the worst lives and leave a relatively select group for the following winter. To his comment that the T-relationship must depend on the scale of temperatures used I would retort that surely that is quite a general point. Any kind of quantitative statement involves the choice of an origin and a unit of measurement.

I do agree with Mr Redington that after saying what I did in the paper about the use of time itself as a factor in analysis it was a pity that in choosing my data for the Appendix I put myself in the position of having to make t my major independent variable.

Mr Kennedy Williams regards my yearning for formal methods of solving algebraic equations as academic. I do not deny that iterative or other approximate methods (often highly ingenious in conception) are extremely useful in practice; but their application can be a very boring business. There are worse hobbies, I think, than trying to preserve in our practical work some flavour of the elegance which—to my mind, at any rate—is one of the most captivating characteristics of mathematical analysis.

My description of what has been called in the discussion 'the cluster method' was intended to be quite non-committal (see paragraph 42); but having regard to the very substantial labour which more orthodox methods of multivariate analysis inevitably involve, I am not sure that it would be quite fair to dismiss the method solely on the ground that it may require an astronomical number of computing operations. The preparation of Table 4 of the Appendix (a very diminutive example, it is true) was not a stupendous effort; with an orderly arrangement of working columns, a table of logarithms and a table of square roots, the 35 sets of values can be obtained in 3 or 4 hours.

It is, of course, inherent in the method that the numerical values selected after consideration of the results adjacent to the mode may not exactly reproduce any one member of the statistical series which represents the dependent variable. What is more to the point, I think, is that the distribution of the $\binom{n}{r}$ solutions may be such that the position of the mode is far less obvious than it is in Table 4; or, conceivably, there may be several modes! At the same time, I cannot help feeling that the method is of some interest in that it shows, before 'expected' values have been computed, the varying degrees in which individual members of the series support, or discredit, the hypothesis represented by the basic equation. On the whole, I would be inclined to say that the theoretical implications of the method deserve rather more consideration than they received either in the paper or in the discussion; and I would have liked to see someone with more courage than I pursue the suggestion in paragraph 42 about the possibilities of sampling.

Turning to the remarks on the paper itself, I can assure Mr Benjamin that its origin and purpose are accurately described in the first paragraph. It was not my intention to make apologies or excuses for anybody. It may be that actuarial methods would have developed rather differently had the full range of modern statistical technique been available in the early days of our profession; but I venture to think that the two major governing factors would still have been those specified in paragraph 3, and that on the basis of these factors a perfectly natural, not an artificial, contrast can be made between the economic analyst and the compiler of a life office mortality experience. But Mr Benjamin and I agree that both parties should have a part in the further development of statistical theory and method. With reference to another remark by Mr Benjamin, my own view is that the economist has a long way to go before his data reach anything like the standards of precision to which the actuary is accustomed.

Mr Kennedy Williams suggests that a remark in paragraph 27 ought to be qualified somewhat in recognition of the powerful and convenient technique of orthogonal polynomials. But in paragraph 27 I was discussing the multivariate case in its most general form, without reference to any of the conditions specified by Mr Kennedy Williams as prerequisite to the employment of the orthogonal polynomials. Where more than one independent variable is involved or the given values are not equidistant, the importation of the orthogonal device, or something akin to it, into the fitting process must, in the nature of things, be a more complicated business. Indeed, I am not aware that a method for the multivariate case has yet been worked out.

I agree with the opener that, having embarked on a discussion of least squares procedure, I might well have said something on the subject of maximum likelihood. My excuse for this omission can only be that I was not attempting to produce a pocket text-book.

I was very interested in Mr Barnett's description of his experiments in fitting the Makeham formula. I did not intend to suggest that what I described in paragraph 30 as the commonly accepted practice is the only method.

The primary object of the paper was to consider differences of technique, rather than differences of function, between the actuary *qua* constructor of life office mortality tables and, say, the economist. Hence, while differences of function were mentioned as one of the probable reasons for differences of technique, no attempt was made to contrast the responsibilities of the one operator with those of the other. Mr Haycocks and Mr Robarts contributed some remarks which have a bearing on this subject. My personal view is that, outside the actuarial field, it is far from easy to say exactly where the function of the statistical analyst begins and where it ends; I can only suggest that he ought to be given the fullest opportunity to contribute both to the planning of the investigation and to the interpretation of the results. In the sphere of life contingencies, it seems to me that the actuary supplies both the doctrine and the technique; and I agree entirely with the view that the final stage—interpretation—is at least as important as anything which precedes it.

In conclusion, I should like to thank all those present at the meeting for their reception of the paper and to say, in particular, how much I appreciated both the kindliness and the wisdom of Mr Redington's remarks in closing the discussion. I should like also to associate myself whole-heartedly with the President's expression of regret at the unavoidable absence of Mr Stone.