JIA 115 (1988) 693-698

On 28 June 1988 a discussion took place on *C.M.I.R.* 9 and *On Graduation by Mathematical Formula*, by D. O. FORFAR, B.A., F.F.A., J. J. MCCUTCHEON, M.A., Ph.D, F.F.A. and A. D. WILKIE, M.A., F.F.A., F.I.A., *J.I.A.* **115,** 1.

GRADUATION BY MATHEMATICAL FORMULA

This paper is about graduation by mathematical formula. There are, of course, other methods of graduation, but, quite deliberately, we have excluded them. Some of you may see this omission as a failing on the authors' part. In our view, however, to have discussed other methods of graduation to anything like the same depth would have made the length of our paper unacceptable. We are enthusiastic advocates of the method of graduation by mathematical formula. It is of extremely wide, if not almost universal, application. One strong motivation for our paper was the thought that we might try to reconcile mainstream modern statistical concepts with ideas and techniques developed by actuaries over many years. We hoped to show how past work by the profession fits in, or may be reconciled, with current statistical theory and at the same time to develop new techniques for future applications.

Throughout we consider three types of mortality rate—the traditional q_x, (which is often called the 'initial' rate), the central death rate m_x, and the force of mortality (or 'instantaneous' death rate) μ_x. All three measures of mortality have played significant roles in the historical development of actuarial science. Over the years, however, q_x and μ_x have played the major roles and q_x has been the basic function for most recently published tables. However there are strong theoretical arguments for considering the force of mortality as the basic building-block. It is worth noting that modern computers enable very precise estimates to be made of the exact exposures to risk theoretically required for dealing with μ_x. These (central) exposures may be used directly for graduations of q_x without the need to convert the central exposures to initial form by some approximation. Moreover many approximations used in the past have limited theoretical justification.

At the recent Faculty discussion of this paper some criticism was made of the fact that no mention was made of available statistical packages, such as GLIM and GENSTAT. Although certain restricted forms of the GM and LGM formulae come within the class of generalized linear models covered by GLIM, this is not true for the full set of GM or LGM families. In theory GENSTAT is a tool to tackle most of the graduations discussed. In practice, however, this does not yet seem to be the case; some of my colleagues have encountered technical problems trying to apply the latest version of GENSTAT to the data of the paper. If these problems can be overcome, and some of them may be inherent in the models we use, then the methods which we advocate and the techniques which we use may become more widely adopted.

The two basic tasks in any graduation by formula are to choose the precise form of formula to be used and to determine the numerical coefficients of the formula adopted. Within the particular context of the GM and LGM families this first problem is simply that of choosing the order (r,s), of the formula to be used. In § 10 we indicate how we choose the 'best' order of formula. It is important to realize that the ideas of this section can be extended to any other classes of formula. Obviously the more complex the formula, the better (in at least some sense) the fit will be for the graduated rates. On the other hand 'over-graduation' of the crude rates can often lead to nonsensical results and will certainly result in much wasted effort. It is important to know when to stop!

We have attempted to reduce the anxiety of readers by deferring to Appendices much of the heavier, but routine, mathematical calculations. The principal results from Appendix 1 are, however, summarized in § 8. Many of the numerical calculations for our work, although conceptually simple, were in practice somewhat complicated and the preliminary computing underlying our paper was done independently on two different machines. The results were then compared. On one machine the optimization was done by 'derivative-free' methods, without using partial derivatives, and on the other use was made of NAG (Numerical Algorithms Group) library routines which require calculation of the relevant partial derivatives. Fortunately the two optimization procedures produced virtually identical results. It is probably fair to say that using partial derivatives, we achieved greater accuracy in the sense that we got closer to the true optimal point, for which there may be underlying theoretical reasons. In order to determine the information matrix, which plays a key role in much of the practical work, the partial derivatives are required and it is preferable to calculate these accurately rather than to estimate them by approximate methods.

ABSTRACT OF THE DISCUSSION

Discussion on paper in JIA, **115,** 1

Mr P. H. Bayliss (opening the discussion): I am a member of the C.M.I. Committee, not involved in its mortality work, but in the work on P.H.I. experience. I was interested in Professor McCutcheon's comments on the difficulty of deriving the information matrix in the way described in the paper. I have done some work on corresponding graduations of P.H.I. experience in this area and have had some difficulty in finding the estimates of the standard errors of my parameter estimates using the information matrix and inverting it and decided to check this by a method of simulation. This may be of interest to those who are trying to apply these ideas in a practical way. Having obtained my graduation results, after having obtained estimates of the parameters. I modelled by simulating a series of experiences, assuming that I had the same exposed to risk as I had been working with in the experience and assuming that the population parameters were the same as the graduation parameters. For each cell of the experience (in mortality terms that would mean each age of exposure), I simulated the actual number of decrements by assuming a Poisson distribution of the numbers of decrements cell by cell, built up a simulated experience for the whole range of the ages, and, in the case of sickness experience, for the sickness durations.

I re-graduated that simulated experience, using exactly the same formula and methods as were used in the previous graduation, and obtained a set of new parameter estimates. I then repeated this a sufficiently large number of times to enable me to calculate the standard errors, and if required the covariances between the new parameter estimates. On checking back I found that the results were virtually identical to those that I had originally obtained by using the information matrix and inverting it. Thus I had found a method of by-passing the information matrix procedure by simply simulating a set of experiences and calculating the standard errors of the parameter estimates in a practical way.

Mr C. D. Daykin: Most actuaries do not have to fit formulae to mortality data to the extent that the C.M.I. Committee have to. For example many people working in pension funds have fairly small data sets available to them, and it would be inappropriate to use curve-fitting techniques in order to produce a set of mortality rates. Nevertheless, there are applications where a good mortality table is necessary, particularly in the life assurance field, and we rely very heavily on the C.M.I. Committee to produce suitable tables for that purpose.

I am concerned in the graduation of the English and Scottish life tables based on population data. I have been reflecting on the work described here and comparing it with that used in preparing life tables based on the three years around the 1981 census. We did not use a mathematical formula approach, as had been used in many previous English Life Tables, but in this and the previous set of English Life Tables we used the method of cubic splines. Perhaps the rationale is that there is an even greater data set available than in the C.M.I. experience. The total population is being dealt with and there is an extremely large number of deaths. The aim is to fit the data closely, much more so than for the C.M.I. data where there are quite a number of outliers, particularly at the younger and older ages.

I was interested in the emphasis which was placed in the paper on the determination of confidence intervals and in the use of sheaves, simulations and quantile plots which illustrate the extent to which there is real confidence in the curve over certain parts of its existence and, at the upper and lower ends particularly, considerable uncertainty. That seems to be a useful tool which helps us to see the extent of the range of the data for which there is good justification for the formula being adopted, and the areas at the upper and lower end, where the data are not pointing to a specific formula. That needs to be borne in mind by practitioners using the tables at the extremities.

Mr R. H. Plumb: There have been for the last 100–200 years mortality tables which are fairly smooth in lay-man's terms rather than in mathematical terms. We are now entering an era where there may be for some time quite violent fluctuations amongst the younger ages, in particular in male assured lives mortality tables, and probably eventually in female mortality tables. How does this work reflect on the likely mortality tables which are going to occur? Will we find ourselves having to use blending techniques between two or three different age-groups of the population simply because of the different types of diseases and viruses affecting different sections of the population? I am concerned that this could be a major problem for us. I wonder whether we have been in a very fortunate position in mortality terms for some time in that we have had a very smooth table with which to operate.

Mr M. H. Westley: We could make more use of the population mortality data that we have as prior knowledge for deriving actuarial tables. Our assured lives and others are subsets of the population. Mr Daykin has explained that it is possible to graduate the population tables much more finely than the assured lives tables because of the larger volume of data. Although the official graduation of the latest English Life Tables did not use a formula, an alternative graduation was carried out by Forfar & Smith (*T.F.A.* **40,** 270) which used a formula based on Gompertz law $q_x/(1-2q_x)$.

We need to keep Gompertz law always in mind and if necessary keep changing the function we are trying to graduate with it to improve the fit at old ages, as the English Life Table graduation leads to an asymptotic value of one half for q_x. Would it not be a reasonable idea to use Gompertz law for graduating our mortality data, but choose the function to be graduated as $q_x/(1-\alpha q_x)$, where α is a parameter chosen to give the best fit, and reject the graduation altogether if it leads to an asymptotic value for q_x which is outside the expected range of 0 to 1, say?

We can use population mortality as a basis for an actuarial graduation by choosing as the base for our actuarial formula for assured lives a formula that seems to fit the population, like the Forfar and Smith graduation of E.L.T. 14. That graduation formula involves several extra terms apart from Gompertz law, one of which is a hump in the mortality at young adult ages. We suspect that AIDS will cause the hump at young adult ages, particularly in the male table, to grow and change in shape over the next decade or two, but it would still be useful to try and represent forecast AIDS mortality as a term in that graduation. In AIDS Bulletin No. 2, the Institute working party have produced some forecast additional mortality rates for AIDS deaths which do not quite fit the shape of the hump in the Forfar & Smith graduation of E.L.T. 14. I should like to see some efforts made to produce a new graduation which allows for the future effect of AIDS by the addition of a suitable time-dependent term.

Professor J. J. McCutcheon, F.F.A.: Mr Westley and Mr Daykin have both spoken about population mortality. In some sense it is always a surprise to me that it is necessary to do any graduation. The population comprises, after all, a fairly large body of people! Intuitively I would have expected rather smooth rates of mortality for the whole population, yet that does not happen. There are strange things, for example with the crude rates of mortality at the older ages. Here it is presumably reasonable to assume that the rate of mortality increases as age increases, yet sometimes the crude rate at one age is less than that at a younger age.

The spline technique has proved useful in constructing the population mortality tables associated with the 1971 and 1981 censuses. This replaced a 7-term formula, which had a rather bell-shaped logistic term, used to produce E.L.T. 11 and E.L.T. 12, and which was found to be no longer appropriate when the 1971 and 1981 census data were examined. The Government Actuary's Department may well look at the possibility of whether the formulae of Heligman & Pollard or of Forfar & Smith could be used the next time. It becomes a balance, because you may need to have a formula with so many terms that it is not worthwhile.

The suggestion was made that everything might be done by formula, for instance $\log(q/(1-q)) + \sin(q/(1+q))$! I am not sure that this would always be very practicable.

Professor A. D. Wilkie: The AIDS model that has been used by the Institute working party is based on a more elaborate pattern than simply transition from live to dead, which is the basis of an ordinary mortality table. Although the model is much more elaborate, the forces of mortality, forces of infection, forces of becoming sick, which were called transition intensities, are assumed to be smooth. For the purpose of the model they have to be clearly defined functions of some kind for each particular age, calendar year and duration. The fact that the structure is more complicated and that the intensities may vary by calendar year produces a table of resulting mortality rates which is nevertheless a fairly smooth two-dimensional pattern rather than the smooth one-dimensional pattern that ordinary mortality rates are usually assumed to be. The graduations of a particular quadrennial mortality experience, like that of 1979–82, are looking at one cross-section of a two-dimensional mortality surface. It would be possible, and I do not know if anybody has ever done this, to take the mortality experience of those people who were born, say, in 1870, most of whom should by now be dead and follow them throughout their entire lifetimes to produce a generation or cohort mortality table for their whole lifetimes. This would be expected to have been reasonably smooth, with possible interruptions at the time of wars and epidemics. There are probably greater fluctuations by calendar year than by age in the underlying rates.

It is also possible to fit two-dimensional surfaces to mortality data if we have enough data like that with cohorts born in each successive calendar year. More recent calendar years produce only a triangular table because there is no experience yet of the future. It is the same problem as that in general insurance of triangular run-off tables.

The AIDS model also produces a triangular table, but it is a forecast table. It is filling in the other half of the run-off table, also based on mortality by calendar year as well as by age. The problem with AIDS is that it is not possible to express the mortality in a single mortality table. It will vary so much by calendar year as well as by age that there is no way in which the extra mortality due to AIDS can satisfactorily be represented by a single mortality table such as we have been using up to now. It would not matter if you tried to fit it with a more elaborate formula. It would need to be a different formula for each calendar year of birth, which in effect is what the AIDS working party has produced and has published specimen rates from.

The two-dimensional pattern has been used, by Bayliss & Waters for the P.H.I. sub-committee of the C.M.I. Committee in dealing with rates of recovery from sickness and of death during sickness where a two-dimensional pattern is appropriate, based on age and duration of sickness, although not a three-dimensional one bringing in calendar years as well.

Professor S. Haberman (closing the discussion): I think it is a pity that the authors have set themselves the rather restricted terms of reference of considering what is traditionally called graduation by mathematical formula. Some of the comments that have been made so far suggested that, for small data sets, these techniques are perhaps like a sledgehammer cracking a nut.

There are other data sets where there might be a set of mortality rates that is changing dramatically by calendar year, perhaps with several humps. This might be as a result of AIDS mortality, although these comments do not correspond to any Bulletins that the AIDS working party is currently involved in. Other methods of graduation which do not require us to produce a complicated mathematical formula might be suitable for such situations. Perhaps a non-parametic method should be considered, one that attempts to smooth the given crude mortality rates rather than fitting a curve to them. As an example I have recently been applying kernel methods to mortality rates (e.g. *J.I.A.* **110,** 135 and *J.I.A.* **114,** 339).

The authors state in § 1 that they "have attempted to reconcile the traditional methods used by actuaries . . . with the more modern work of statisticians in the fields of survival data and life testing". I agree that they have managed to reconcile our methods with the theoretical work of statisticians, bringing in, for example, the information matrix, but I do not think that they have reconciled our work with that of statisticians who are working in the analysis of survival data and life testing.

Why do I think that? There is a brief mention of the work of Cox & Oakes. Cox, in 1972, invented what is now called survival analysis which is a type of graduation; a type of smoothing of data where allowance is made for the effects of other information or covariables. For example, if considering a group of impaired lives the covariable information might be blood pressure, or some other information on medical status. An attempt is being made to represent the hazard rate, as the statisticians call it, or the force of mortality, as we call it, in some sort of regression on these co-variables.

Another area in which we have statisticians working in a parallel way to actuaries is in exposed to risk. There are statisticians who have approached the estimation of mortality rates using maximum likelihood methods and have shown, for example, how accurate is the traditional actuarial approach based on exposed to risk. For example Elandt-Johnson & Johnson have shown that our methods are quite robust compared to more sophisticated statistical approaches. We should be aware of what other scientists are doing in areas that we perhaps regard as our own province.

It is worth noting that in § 13, where two experiences are compared, what is happening is that two sets of rates are compared in their entirety. Graphs of two sets of q_x or μ_x are not being drawn, looked at for the widest gap between the two and then tested to see if that is significant. That is cheating. The rates are compared in their entirety.

When the data set is small, unlike the C.M.I. data, there are other ways of approach which do not involve the use of the information matrix, for example non-parametric ways of counting the numbers of deaths by time in the two or more groups under consideration.

In § 9 there is a full discussion of various tests of a graduation. I do not think that this is a deficiency on the part of the authors, but rather of the text books that we and our students use. Each test is aiming to test a particular hypothesis. Of course, there are dangers if there are many tests, each trying to test the same hypothesis. Eventually a significant result will be achieved. What does that mean? Such caveats on the dangers of multiple hypothesis testing do not appear in any of our text books. They are implicit, I think, in this paper.

I am not sure that it is surprising that mortality rates level off as age advances, as shown in § 17.8, considering the ultimate experience of male insured lives. This is the underlying thesis that Beard suggested many years ago. It is likely that a cohort of people who are ageing over time constitutes a mixed population. Some are stronger than others. The weaker people die and the slope of the curve of age-specific mortality rates decreases as age advances. There is a point of inflection in this curve, reflecting an improvement in the relative strength of the population with age.

Such a picture is well-documented in a range of mortality data: in America the mortality rates for non whites at very high ages fall below the mortality rates for whites; the mortality rates for males overtake the mortality rates for females. These cross-overs can possibly be explained by the population under analysis being heterogeneous. Even more so with C.M.I. data, where the population is mixed by way of the life office concerned and by way of the method of sale, e.g. broker or direct sales.

How do the authors see this subject developing in the future?

Professor J. J. McCutcheon, F.F.A. (replying to the discussion): The question was raised as to what happens as you get older. The Registrar-General, Mr Thatcher, wrote a paper on centenarians (*J.I.A.* **114,** 327). One of the professors at the University of York, has been invited by Mr Thatcher to do further joint work on the mortality of old people.

There were areas that we deliberately chose not to explore. We shall think carefully about what Professor Haberman said.