Bart Holland

On the Adequacy of Mitra's Model of the Life Table : A Technical Note

A variety of mathematical functions have been used as approximations for the purpose of modelling the life table function. Gompertz was an early historical figure whose contributions mark the starting point for much mathematical work in the area. His relatively simple model has given way to models of greater complexity with more numerous parameters, and these models provide ever-closer approximations to observed life table data. Makeham 1960), Perks (1932), Pollard (1979), and Siler (1983) have all contributed to this progression. Mitra (1983) has recently provided a most interesting mathematical treatment of the life table resulting in a model which, logic suggests, will recapitulate quite closely the patterns of observed life tables. This model has the advantage of parsimony in that relatively few parameters are required. However, Mitra's case is not as compelling as it might otherwise be, because of the way the model is tested.

Mitra's model (1983) is tested by generating parameters which permit the model to replicate the Coale-Demeny life tables (1983). The Coale-Demeny tables are an aggregate representation of recorded mortality patterns and are classified by patterns and levels. The question is whether there is good agreement between *l(x)* values from the Mitra model curve and the Coale-Demeny curves: If goodness-of-fit is substantial, then the mathematical formula captures the observed patterns of mortality and can be a useful tool.

We are presented with multiple correlation coefficients, R^2, as a measure of goodness-of-fit of Mitra's model. Goodness-of-fit, in the sense of agreement of two sets of values, in not correctly measured by correlation coefficients. A simple example will clarify this point. Suppose we have two sets of values, *x*

and *y*, and that *x* ranges from 1 to 100. If $y = x + 99$, then the two measures will be perfectly correlated, but there will always be a big difference between them—indeed, a difference as large as the entire range of *x*. The relation between *x* and *y* is nonetheless a perfect linear correlation, because a plot of *x* and *y* lies on a perfectly straight line. Thus, we see that the correlation coefficient measures linear association or co-variation, and not identity, of two sets of values.

By extension, the multiple correlation coefficient R^2 measures co-variation of a variable with a function which has several parameters, in a relationship obtained by multiple regression, and again the multiple correlation coefficient does not measure identity or agreement. Moreover, the error structure of least squares regression (by which these R^2 values were presumably estimated) does not hold for *l(x)* curves, because the values along an *l(x)* curve at different levels of *x* are not independent. Such curves are constrained to a monotonic decline and survivors are represented at more than one point. Hence the R^2 values do not even accurately reflect the proportion of variability explained by linear association. Finally, the goodness-of-fit of the model is not compared to the adequacy of any other model. It is impossible to judge whether the model is any improvement over other mathematical formulations of the life table.

It may be objected that correlation has been used in the construction of some generally accepted life tables, such as the Coale-Demeny tables themselves (1983). In constructing the Coale-Demeny tables, the question arose as to whether a high or low value of *l(x)* early in life was correlated with high or low values at later ages. In fact, a strong positive correlation of *l(x)* values does hold in general over the course of the life table, which simply means that a high mortality population tends to have high mortality at every age, and that the converse is true as well. The quantification of this finding was important to Coale and Demeny, but co-variation and not identity was a matter of interest.

What could be done to validate Mitra's model more effectively? One approach to testing would be to examine *l(x)* curves for patterns of deviation from expectation. This method is well-described by Smith (1983). Suppose we have a set of data, and a model curve which we say fits the data well. If observations deviate from the model curve in a systematic fashion—say, the model curve is too high in one part and too low in another—the model probably does not fit well. If there is no systematic pattern of deviation then the deviation is probably due to random variability, and the model is probably good. Durbin-Watson statistics quantify the degree to which deviations are systematic or not. Smith (1983) provides instructive advice on using computer packages to generate probability levels for the goodness-of-fit of model life tables using Durbin-Watson statistics, and these probability levels can be used to test statistical hypotheses about particular models.

To test whether the model reproduced a specific reliably-recorded life table (rather than an aggregate table like the Coale-Demeny system), we might use a chi-square test of goodness-of-fit. The observed distribution of deaths by age would then be compared with the corresponding distribution generated by the model. Note that we test differences in observed and expected numbers of deaths, not *l(x)* values. The chi-square test requires separate observations in each age interval, but in *l(x)* curves survivors are represented at more than one point. The mechanical details of applying this test are found in Lee (1980). The null hypothesis of the statistical test is that there is no difference between the observed and expected distributions of deaths, i.e. that the data do not deviate (except for sampling error) from what the model predicts.

As to Mitra's model, we are only presented with R_2 as a measure of goodness-of-fit; we do not have tables showing the values generated by the model using his selected parameters, together with the corresponding Coale-Demeny values which he is fitting. A graphical presentation, which would provide a basis for an informal visual assessment of fit, is also absent. Thus, from the information as presented.we can neither confirm nor reject the hypothesis that this model provides an excellent fit to aggregate observational data. The point is technical, but not trivial : in modelling the life table curve, the goal is to replicate extremely closely the observed patterns of mortality, for such a mathematical formulation would have many practical uses. A particular model is useful only insofar as it is judged to meet this goal. The degree to which Mitra's model replicates observed mortality patterns is still an open question.

References

1. Coale, A. J., Demeny, P. and Vaughan, B., 1983, Regional Model Life Tables and Stable Population, Academic Press.
2. Lee, E. T., 1980, Statistical Methods for Survival Data Analysis, Lifetime Learning Publications, Belmont, California.
3. Makeham, W. M., 1960, On the law of mortality, Journal of the Institute of Actuaries, **13,** 325-358.
4. Mitra, S., 1983. Modelling life table functions, Demography India, 12 (1), 115-121.
5. Perks, W., 1932, On some experiments in the graduation of mortality statistics, Journal of the Institute of Actuaries, 63, 12.
6. Pollard, J. H., 1979, Factors affecting mortality and the length of life. In : Population Science in the Service of Mankind, Conference on Science in the Service of Life. Vienna, 1979, International Union for the Scientific Study of Population, Liege, 59-72.
7. Siler; W., 1983, Parameters of mortality in human populations with widely varying life spans, Statistics in Medicine, 2 (3), 373-380.
8. Smith, D. P., 1983, Durbin-Watson statistics for model life tables, Asian and Pacific Census Forum, 9 (4), 7-9.