Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Simplified Azbel Model for Fitting Mortality Tables

Cvetko J. Andreeski*, Pandian Vasant**

*Cvetko Andreeski, Faculty of Tourism and Hospitality – Ohrid, Macedonia
E-mail: cvetko.andreeski@uklo.edu.mk
**Pandian M. Vasant, Electrical & Electronic Engineering, Universiti Teknologi Petronas, Kingdom of Malaysia.
E-mail: pandian_m@petronas.com.my, pvasant@gmail.com.

Abstract: Since the 1825 there is Gompertz mortality function for estimation of mortality data. All the attemts of finding apropriate mortality function lead to functions that are not easy to fit with the actual mortality data. The most recent reasearches (Azbel) offer simple function for fit mortality tables. Common characteristic of all these functions are difficulty of fiting data. We need to use numerical approach to find the appropriate values of function parameters to fit the function to the mortality data. In this paper we suggest simplified Azbels model which offers more simple technique of fitting data, the well known maximum likelihood estimator. This function is simplifyed to one parametar function and offers fitting data with one equation.

1. INTRODUCTION

GOMPERTZ (1825) observed an approximate exponential increase in death (mortality) rates with age. A number of authors (Gavrilov, Loschky, Strehler, Mildvan) have put forward biological theories to explain this relation. These theories are summarized in Mildvan and Strehler et al. 1960. Delineating the causes of death and age ranges that give rise to mortality rates adhering to the Gompertz relation should prove useful in sharpening theories of the type mentioned above. Further, the exponential relation between mortality rate and age is frequently taken as standard. A substantial departure from this relation is then taken to suggest that, for a subset of ages, the population tinder study is encountering an unusual risk or stress.

There are many function that aproximates mortality tables. But they are too complex or inapropriate to use them to get maximum likelihood estimation (MLE) for the parameters.

If we take a look at Gompertz mortality function:

$$l(a)=k\cdot g^{c^a} \qquad (1)$$

we can see that it uses 3 parameters. It's shown that functions with more parameters can fit better to the data, but it's very difficult to find optimal values of the parameters. We have the same situation with Gomperz-Makeham mortality function:

$$l(a)=k\cdot s^a\cdot g^{c^a} \qquad (2)$$

because it uses one parameter more than Gompertz function and it's more complex than the Gompertz model.

So, we need to use numerical optimization techniques to find aproximation of the parameter values. Better aproach should be if we find the maximum likelihood estimator for some mortality function and calculate the parameters in analytic form. The objective of this paper is to find maximum likelihood function for estimation of the mortality function.

Maximum likelihood estimation is a totally analytic maximization procedure. It applies to every form of censored or multicensored data, and it is even possible to use the technique across several stress cells and estimate acceleration model parameters at the same time as life distribution parameters. Moreover, MLE's and Likelihood Functions generally have very desirable large sample properties:

- they become unbiased minimum variance estimators as the sample size increases
- they have approximate normal distributions and approximate sample variances that can be calculated and used to generate confidence bounds
- likelihood functions can be used to test hypotheses about models and parameters

We use simplifyed Azbel model to take advantages of MLE of the parameters. This model in most cases as good and even simpler then Gompertz model, with readily interpretable parameters (Zempleny). Results of the fit are given with apropriate statistics compared with other models.

978-1-1234-7890-2/08/$20.00 © 2008 IFAC

10.3182/20080706-5-KR-1001.1907

2. SIMPLIFYING AZBEL MODEL

Azbel model is the simplest model that describes the mortality function and thus best choice for identifying mortality tables. It's given by the following formula (3):

$$q_x = \text{Abexp[b(x-X)]} \qquad (3)$$

Where A, b and X are parameters. This formula can be further simplified by a reparamterization. If we let T=X-(1/b)ln(Ab), we got the formula (4):

$$q_x = \exp(b(X - T)) \qquad (4)$$

Parameters have the following roles: b is the shape parameter and T is the end point as qx is equal to 1 when x is equal to T.

But if we eliminate the parameter A (it gives the amplitude of the function) we should use normalized values of mortality tables in range (0-b). In that case we have simple Azbel model given with the following equation:

$$q_x = \text{bexp[b(x-X)]} \qquad (5)$$

Parameter X has the same role as it has T in equation (4). X is the end point as qx is equal to 1 when x is equal to X.

This model is simple enough, it uses only one parameter and we can find a maximum likelihood estimator (MLE) for this model. We can see that this distribution is similar to the exponential distribution. The equation (5) is one parameter equation and we can find the MLE. So we can proceed to calculation of MLE.

$$L(b) = \prod b e^{b(x-T)} \qquad (6)$$

Using the equation (6) we can find the *l(b)=ln(L(b)):*

$$l(b) = n\ln(b) + b\sum x_i - nbT \qquad (7)$$

If we find the derivation $l'(b)$, we got>

$$l'(b) = n/b + \sum x - nt \qquad (8)$$

Finaly, the maximum likelihood estimator for this model is given with the equation (9).

$$b = \frac{n}{nT - \sum x_i} \qquad (9)$$

3. TESTING THE MODEL

We have tested this model on mortality taken from the Society of Actuaries and we make comparison with the results of other models for fit (Andreeski et al. 2001, Zempleny), as mortality tables of R. Macedonia and Hungary. For the model we have calculated chi square and A/E statistics for testing hipotesis. Results of our testing are compared with similar models calculated with numerical methods.

In the following fig. 1 we have presented graphical fit to the data of mortality rate and survival probability.

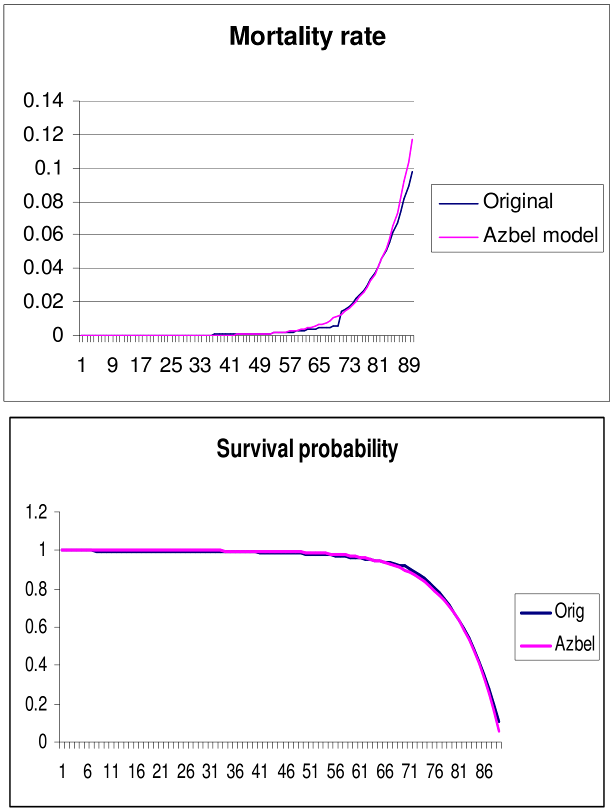

Fig. 1. Graphical presentation of mortality rate and survival probability

Similar to chi-square test is the quadratic deviation or wighted quadratic deviation statistics (QDEV):

$$QDEV = \sum_{i=K}^{N} \frac{T_i (q_{i1} - q_{i0})^2}{365 \cdot q_{i0}} \qquad (10)$$

where Ti is the number of exposures in year i and q1 is the table chosen as a candidate for being the approximator. The formula allows for appropriate choice of the starting and finishing years K and N to be chosen appropriately to the problem we want to solve and to the available data . In formula (1) the weights are chosen as being proportional to the reciprocal of the approximate variance: $\frac{365 \cdot q_{i0}}{T_i}$ of the estimator qi0, ensuring the limit distribution being chi-squared, so a statistical test for the equality of the two tables can be based on the critical values for the chi-squared distribution with N-K+1 degrees of freedom. Besides the fact that this statistic is frequently used, we didn't use this statistic because the number of exposures in year i for our table is 1, so the classical chi-square statistic is enough for testing the fit.

With equation (10) the A/E statistics is defined:

$$A/E = 100 \cdot \frac{\sum_{i=K}^{N} l_{i0} q_{i1}}{\sum_{i=K}^{N} l_{i0} q_{i0}} \qquad (11)$$

This statistics can be found in the paper of Mitchell and McCarthy [9]. In (10) l_{K0} can be chosen as 1, so the summands in the denominator are proportional to the estimated number of deaths in the given year in the population, while the numerator gives the similar quantity, based on the population distribution of the investigated table and the risks defined by the approximating table. This can also be interpreted as a Laspeyres index of the two sets of probabilities q_i, taking weights from the basis population.

Table 1: Table of statistics

Chi square	**Chi square num. method**	**A/E**	**A/E numerical methods**
0.187029 $[69.126]^{0.95}$	0.2232 Maced.	103.5659	95.64 Maced.
	0.54932 Hung.		111.09 Hung.

From the table we can see that chi sqaure value is very low (critical value for 89 degrees of freedom is 69.126), which gives us information that the model gives good git to the data. Ideal value of A/E statististics is 100. Our A/E statistics shows close value to the value of 100. We can conclude that the model is valid and it gives good aproximation of the real data. If we compare the value of A/E statustics with parameter calculated with unweighted least square estimates, where the minimization was done by numerical methods we can see that value calculated as MLE gives better fit on data than the numerical methods.

The interested reader may find further test statistics in the well-known book of (Benjamin and Pollard et al. 1993).

On the following figure 2. we show the comparative graphical presentation of original data, Azbel model fitted with MLE and Azbel model fitted with numerical estimation.

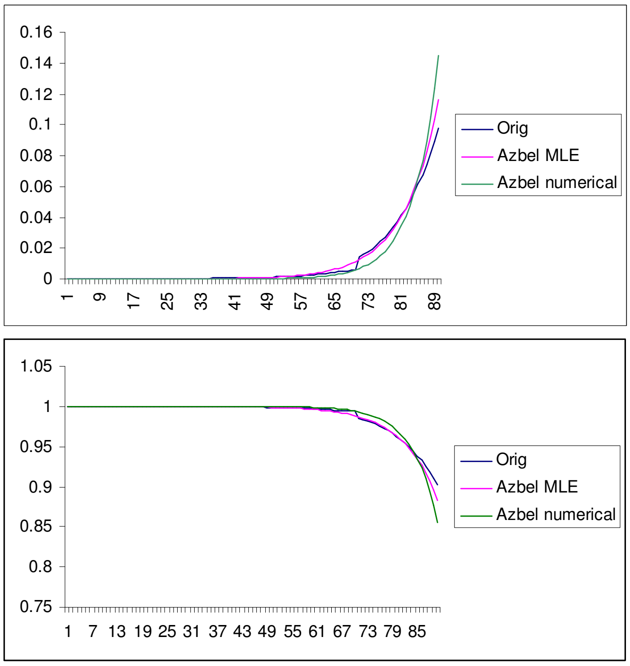

Fig. 2. Graphical presentation of mortality rate and survival probability with Azbel models MLE and numerical

4. ADVANTAGES OF USING MLE

We can calculate model parameters with many various techniques: method of moments, unweighted and weighted least square methods, and MLE . What are the advantages of using maximum likelihood estimator for calculating function parameters. There are several advantages we should take in account:

- For a fixed set of data and underlying probability model, maximum likelihood picks the values of the model parameters that make the data "more likely" than any other values of the parameters would make them. We cannot find better estimator for the data than the MLE.
- For the function with one unknown parameter, using MLE gives also advantage of fast calculation of the parameter. The parameter value is given with one analytic equation (9). We don't need complex models of complex methods for calculations.
- If we need different functions with different values of the parameters we can easily calculate all this functions very fast. One example of such a need is forecasting future values of mortality in some country or countries (Zempléni, Andreeski). This is almost always the case in insurance. In the new EU member states, where recently there were quick changes in mortality patterns we may expect the changes to continue and we need to make forecast of the mortality. Annuities are also calculated on the bases of mortality forecast. In order to have more accurate

calculations we need to have more accurate mortality forecast.

5. CONCLUSIONS

Maximum likelihood function gives best estimation of any other model of parameter estimation. Function that models the mortality are too complex or inapropriate for calculation of maximum likelihood function in analytic form. They all have many parameters that should be taken under consideration. Simplest model for modeling mortality is the Azbel model that has only two parameters in it's basic form. This model with calculation of data normalization can be used in calculation of maximum likelihood function. All the models of mortality function are intended for modeling with numerical optimization techniques. Even with numerical optimization techniques it's almost impossible to find optimal values of all different parameters of the model.

Azbel model fitted with MLE gives good results, even better then the estimations using unweighted and weighted least square methods with numerical models. If we see the chi square (or quadratic sqaure statistics) and A/E statistics of the compared mortality tables we can conclude that simplified Azbel model gives better results in fitting data then other models and techniques of estimation. On the end we should emphasize again some of the advantages of MLE besides the best fit of the parameters.

MLE has many optimal properties in estimation: sufficiency (complete information about the parameter of interest contained in its MLE estimator); consistency (true parameter value that generated the data recovered asymptotically, i.e. for data of sufficiently large samples); efficiency (lowest-possible variance of parameter estimates achieved asymptotically); and parameterization invariance (same MLE solution obtained independent of the parametrization used). In contrast, no such things can be said about least square estimator.

REFERENCES

András Zempléni, Forecasting and simulating mortality Tables, www.math.elte.hu/~zempleni/morttabcikk_0324.doc

Azbel, M.Y. Empirical laws of survival and evolution Proc. Natl. Acad. Sci. 1999

Andreeski, C.J. "Analysis of Economic Time Series: Algorithms and Application Software for Model Developement." ASE-ETF Technical Report ETS-01/2001. SS Cyril and Methodius University, 2001.

Benjamin, B. and J. H. Pollard. 1993. The Analysis of Mortality and other Actuarial Statistics, Oxford: The Institute of Actuaries and Faculty of Actuaries.

Box, George E.P. *Time Series Analysis - Forecasting and Control*. New Jersey: Prentice-Hall, 1994.

Chatfield, C. *The Analysis of Time Series - An Introduction* (Fourth edition). London etc.: Chapman & Hall, 1989.

Dr. Leonid A. Gavrilov, Ph.D., Dr. Natalia S. Gavrilova, Ph.D. Biodemographic Reliability Theory of Ageing and Longevity, Center on Aging NORC and The University of Chicago Chicago, Illinois, USA, 2007

Gujarati, Domodar N. *Basic Econometrics* (Third edition). New York etc.: McGraw-Hill, 1995.

Loschky J. David, Economic Change, Mortality and Malthusian Theory, *Population Studies*, Vol. 30, No. 3 (Nov., 1976), pp. 439-452, 1976

McCarthy, David G. and Olivia S. Mitchell, 2001. "Estimating International Adverse Selection in Annuities", PRC Working Paper, 2001, University of Pennsylvania, Philadelphia

Miklós Arató, Dávid Bozsó, Péter Elek, András Zempléni: Forecasting and Simulating Mortality Tables. 2006

Mills, Terence C. *The Econometric Modelling of Financial Time Series*. Cambridge: Cambridge University Press, 1993.

Strehler, B.L. and Mildvan, A.S. (1960) General theory of mortality and aging. *Science 132*, 14.