ABSTRACT

Actuaries are often in the position of having insufficient data at advanced ages for their purposes. Reliable mortality data tends to stop at ages well below the ones to which actuaries need to make calculations or projections. This paper describes how to use two different classes of mortality-projection model to extrapolate mortality rates by age as well as projecting forward in time. The first model used is the 2D age-period model from Currie, Durban and Eilers (2004), which is used here in a new application for extrapolating in the age direction. The second model is a new proposal from Currie (2010), which in turn is a generalization of the model from Cairns, Blake and Dowd (2006).

KEYWORDS

mortality projections; smoothing; longevity risk; P-spline models, Cairns-Blake-Dowd models.

CONTACT ADDRESS

Stephen Richards, 4 Caledonian Place, Edinburgh, EH11 2AS, United Kingdom; Telephone: +44 (0)131 315 4470; Email: stephen@longevitas.co.uk; Web: www.longevitas.co.uk

Iain Currie, School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, United Kingdom, EH14 4AS.

1. Introduction

Actuaries often find themselves without data at advanced ages for modelling and projecting mortality rates. One example is the ONS mortality data for England and Wales; data are available for males and females from 1961 onwards, but they stop at age 104. Another example is the CMI “assured lives” data set, where the quality of that data is questionable beyond age 95 — see Figure 1. In either case, actuaries need a means of extrapolating mortality rates from the last available reliable age up to the last age required for calculation purposes. This article describes how this can be done with two quite separate classes of model which are ordinarily used for mortality projections.

© Stephen Richards and Iain Currie. All rights reserved.
Figure 1. Logarithm of crude force of mortality for CMI data at ages 50–100 aggregated over the years 2001–2005. The data above age 95 are unreliable and cannot be included in any projection model. However, actuarial calculations for annuities and pensions typically require mortality rates up age 120, so some form of extrapolation is required from age 95 to age 120.

A related issue lies with mortality projections themselves. The advent of Solvency II has placed an emphasis on “probability distribution forecasts” (European Parliament and European Council, 2009), which suggests that stochastic methods are required for all aspects of portfolio valuation. In the case of longevity risk in pensions and annuity business, this means building a stochastic projection model for future mortality rates. In addition to needing extrapolated mortality rates at advanced ages, actuaries also require corresponding measures of uncertainty. The article shows how the methods provide standard errors on the extrapolated rates.

2. Data and Data Preparation

The data used in this paper are the number of deaths aged $x$ last birthday during each calendar year $y$, split by gender. Corresponding mid-year population estimates are also given. The data therefore lend themselves to modelling the force of mortality, $\mu_{x+\frac{1}{2}, y+\frac{1}{2}}$, without further adjustment. We use two such data sets, one provided by the Office of National Statistics (ONS) and one by the CMI.

We use ONS data for England & Wales for the calendar years 1961–2007 inclusive. This particular data set only has death counts and estimated exposures up to age 105, with older ages grouped into a category labelled “105+”. We will work here with the subset of ages 50–104 which is most relevant for modelling post-retirement mortality. ONS death data are provided by the date of registration. In this paper we have used the registration data throughout for consistency. This is the same data set as used in Richards, Kirkby and Currie (2006), but with a smaller age range and some extra years of data.
Extrapolating Mortality Projections by Age

The data set used in Figure 1 is the experience data for assured lives collected by the CMI for calendar years 1947–2005 inclusive. For consistency with the ONS data set, we will use the same age range 50+. Note that the CMI data set is only useful for male lives, as the data are sparse for females.

3. Model structure

The data used in this article lend themselves to modelling the force of mortality, \( \mu_{x+\frac{1}{2},y+\frac{1}{2}} \), without further adjustment. Our model for mortality is that the number of deaths aged \( x \) last birthday in calendar year \( y \), \( D_{x,y} \), follows a Poisson distribution:

\[
D_{x,y} \sim \text{Poisson} \left( E_{x+\frac{1}{2},y+\frac{1}{2}} \times \mu_{x+\frac{1}{2},y+\frac{1}{2}} \right)
\]  

(1)

where \( E_{x+\frac{1}{2},y+\frac{1}{2}} \) is the central population exposed to risk at age \( x \) in the middle of year \( y \) and \( \mu \) is the corresponding force of mortality. Our models will be fitted using the method of maximum likelihood.

All the data here are supplied aggregated, and we will model the mortality of groups. This is in contrast to the models of individual mortality which are used for detailed life-insurer data, as outlined in Richards (2010). Note that both the models for groups in this paper and the individual models of mortality in Richards (2010) are all models for the force of mortality (hazard function), rather than for the mortality rates often used by actuaries. Mortality rates \( (q_x) \) can of course be precisely calculated from knowledge of the force of mortality as follows:

\[
q_{x,y} = 1 - \exp \left( - \int_0^1 \mu_{x+s,y+s} ds \right)
\]  

(2)

or else using the following approximation:

\[
q_{x,y} \approx 1 - \exp \left( -\mu_{x+\frac{1}{2},y+\frac{1}{2}} \right)
\]  

(3)

Figure 2(a) shows the most common situation for a projection model such as those in Currie, Durban and Eilers (2004), Cairns, Blake and Dowd (2006) and many others. We have mortality-experience data covering the cross-hatched area, and we want to project mortality rates into the future, represented by the coloured area. Figure 2(b) shows the situation of an actuary graduating a mortality table. Mortality rates must be graduated (smoothed) at the ages at which data are available, but the demands of actuarial work necessitate the extrapolation of mortality rates by age. This is often done by means of a parametric model (Richards, 2010), which yields fitted mortality rates outside the data area. Extrapolation can be either up or down in age (or both). Figure 2(c) shows the modern situation for an actuary requiring both extrapolated rates in age and projected rates in time, together with measures of uncertainty over those rates.
Figure 2. Three constellations for mortality modelling. The cross-hatched region shows the available data, while the coloured area shows the years and/or ages for which we want to project or extrapolate mortality rates. (a) Projection model — mortality rates are projected in time, (b) Extrapolation model — mortality rates are extrapolated in age, either up to advanced ages or down to early ages (or both), (c) Combined model — model rates are extrapolated in age and projected in time, the subject of this paper.

4. Two-dimensional penalised-spline models

Currie, Durban and Eilers (2004) introduced the idea of projecting mortality rates in time by using a smooth penalty function. The simplest case is of a flexible two-dimensional basis of $B$-splines, where the regression coefficients are smoothed using a penalty function (Eilers and Marx, 1996). Projections in time come from extrapolating the penalty function, which can be done on either an age-year basis or an age-cohort basis (Richards, Kirkby and Currie, 2006). In both cases, the future is treated as missing data and the penalty provides a smooth projection of future mortality rates. Mathematically, the matrix of death counts and exposures is extended to the projection year and filled with dummy data. To avoid this dummy data affecting the model fit, a weight matrix is created which has the value 1 in the region of the data and the value 0 in the non-data region. A classical set-up for a projection model is shown in panel (a) of Figure 2 — the data area is represented by the cross-hatched region, while the non-data area is represented by the coloured region. The basis of $B$-splines would be extended to the projection year, as shown in the lower panel of Figure 3.
Figure 3. Two-dimensional B-spline basis by age and calendar year, showing the extra basis splines for ages for which we have no data (upper panel) and the extra basis splines required for projecting in time (lower panel). In each case the spacing between spline knots is 5 years. The splines used in this paper are cubic splines (de Boor, 2001), but any function with a local peak would suffice.

Figure 4. Logarithm of crude force of mortality for ONS data for males in 2007 at ages 50–104, together with the fitted and extrapolated rates to age 120 under the 2D age-period model (Currie, Durban and Eilers, 2004). As with a projection in time, the uncertainty over the extrapolated rates increases the greater the distance from the actual data, i.e. there is an expanding “funnel of doubt”. In this case the funnel is rather narrow because of the strength of the age signal.
However, ages outside the data range are also a form of missing data. The same 2D models can therefore be used to extrapolate by age in place of projection in time. The \( B \)-spline basis is extended to the maximum age for which rates are required, as shown in the upper panel of Figure 3. The matrices of deaths and exposures are extended in the same manner as is done for future calendar years. This allows the actuary to smoothly extend mortality rates up or down outside the data region. The set-up for this age extrapolation is shown in panel (b) of Figure 2, which is the situation for the creation of a static life table over an age range wider than the available data. Figure 4 shows the output for just such an age-extrapolation exercise for ONS males in calendar year 2007. Similarly to their use for projections in time, the extrapolated rates are accompanied by appropriate standard errors.

For two-dimensional \( P \)-spline models, however, it is also possible to simultaneously extrapolate by age and project in time. The set-up is shown in panel (c) of Figure 2. As before, dummy data are created for the expanded matrices of deaths and exposures, together with the appropriate weight matrix. The basis of \( B \)-splines for age and time are therefore as shown in the upper and lower panels of Figure 3. The result of a simultaneous extrapolation and projection according to the two-dimensional age-period \( P \)-spline model is shown in Figure 5.

![Figure 5](image-url)

Figure 5. Logarithm of crude force of mortality for ONS data for males. Mortality rates are simultaneously extrapolated to age 120 and projected to 2050 using the two-dimensional \( P \)-spline model of Currie, Durban and Eilers (2004).
Figure 5 shows the entire mortality surface after extrapolation and projection. Figure 6 shows the actual mortality rates, together with the projected rates and confidence intervals at four selected ages.

![Graphs showing mortality rates for different ages over time](image)

As described in Djeundje and Currie (2010), the \( P \)-spline model can be fitted with an over-dispersion parameter. This has been done here, with the over-dispersion parameter estimated at \( \psi^2 = 4.31 \).

5. **Cairns-Blake-Dowd model**

Richards and Currie (2009) demonstrated that the choice of projection model can have as big an impact on actuarial calculations as the uncertainty produced within a given model itself. It is therefore useful for actuaries to have other models which are capable of age extrapolation other than those which use penalty functions for projection. One such example is the bivariate time-series model for the force of mortality proposed by Cairns, Blake and Dowd (2006):

\[
\log \mu_{x,y} = \kappa_{0,y} + \kappa_{1,y} (x - \bar{x})
\]  

(4)

where \( x \) represents age and \( \kappa_0 \) and \( \kappa_1 \) form a bivariate time series tracking changes in the level of mortality over time (\( \kappa_0 \)) and changes in the rate of change with age (\( \kappa_1 \)). This model was extended by Currie (2010) as follows:

\[
\log \mu_{x,y} = \kappa_{0,y} + \kappa_{1,y} S(x)
\]  

(5)
where $S(x)$ is any smooth function of age, which includes the straight-line assumption of Gompertz (1825) implicit in Equation 4. Since $S(x)$ is a smooth function of $x$, it is capable of extrapolating mortality rates by age. Currie (2010) noted that where $S(x)$ is formed from penalized splines, the resulting fit to the observed data was much superior to the Gompertz assumption where $S(x)$ was a simple straight line. The parameters resulting from this smoothed $S(x)$ version of the Cairns-Blake-Dowd model are shown in Figure 7.

![Figure 7](image_url)

Figure 7. Parameters for the Cairns-Blake-Dowd model number 5, as modified by Currie (2010). The function $S(x)$ is smoothed by splines with a five-year knot spacing, which enables age extrapolation as in the 2D $P$-spline models. The parameters $\kappa_0$ and $\kappa_1$ form a bivariate time series for projecting future rates, which is a quite different structure to the penalty projections of the $P$-spline models.

Unlike the $P$-spline models, it is not necessary to make any kind of direct allowance for over-dispersion. Instead, the over-dispersion will be picked up as extra volatility in the time-series process. The resulting two-dimensional mortality surface is different under the Cairns-Blake-Dowd model, as shown in Figure 8. Mortality rates in the data region exhibit ridges due to the fitted period effects, while there is evidence of greater curvature with age than in the 2D age-period model shown in Figure 5.
Figure 8. Logarithm of crude force of mortality for ONS data for males. Mortality rates are extrapolated to age 120 using the $S(x)$ function, and projected to 2050 using the bivariate time series for $\kappa_0$ and $\kappa_1$. The fitted rates have period-effect ridges, unlike Figure 5, while the fitted and projected rates have more curvature with age than in Figure 5.
6. Other models

Not every projection model is capable of extrapolation by age. To illustrate why, consider the structure for $\mu_{x,y}$ from Lee and Carter (1992):

$$\log \mu_{x,y} = \alpha_x + \beta_x \kappa_y$$

(6)

where $\alpha_x$ is the mortality effect by age $x$, $\kappa_y$ is a time-dependent effect and $\beta_x$ is an age modulation of the time factor, $\kappa_y$. For simplicity we have left out the half-year adjustments, but the model will actually apply at ages 50.5, 51.5, ... and at the mid-points of each calendar year $y$.

Figure 10 shows the fitted parameters according to the Lee-Carter model. Since the Lee-Carter model uses point estimates for $\alpha_x$ and $\beta_x$, these values are undefined at ages outside the range of the data. The Lee-Carter model cannot therefore be used for age extrapolation like the 2D $P$-spline or CBD5 models. Only models which have a smooth functional form for the age-related component of mortality can be used to extrapolate mortality rates by age.
Figure 10. Parameters for the Lee-Carter model resulting from fitting to ONS data for males aged 50–104. The identifiability constraints applied are \( \sum \kappa_y = 0 \) and \( \sum \kappa_y^2 = 1 \). The fact that the model uses point estimates for \( \alpha_x \) and \( \beta_x \) means that these parameters are undefined outside the data range and therefore the model cannot be directly used ‘as is’ for age extrapolation.

7. Conclusions

This paper shows that there is a number of mortality-projection models available for extrapolating mortality rates by age as well as projecting in time. These models can be used to extend mortality tables beyond the ages at which the actuary has data. For Solvency II and other purposes requiring stochastic mortality projections, these models allow the sensible extrapolation of rates at the highest ages along with consistent future projections of those rates.

Disclosure of interests

Stephen Richards is a director of Longevitas Ltd and has a financial interest in the Projections Toolkit modelling software used in the work in this paper.

Acknowledgements

The authors thank Adrian Gallop of the GAD for the population experience data. Any errors or omissions remain the sole responsibility of the authors.

Model fitting was done using the Projections Toolkit, while R was used for all graphs. Typesetting was done in pdfTEX.
References


