Overview of Lee-Carter mortality modeling

Demography 215

October 19, 2010

Abstract

The Lee-Carter model is actually a collection of several techniques and ideas which combine to produce a suite of powerful methods allowing data summarization and population forecasting. A simple model converts an array of centered age-specific rates into a relational model involving an age profile $b(x)$ scaled by a simple time index $k(t)$. The historical variation in the time index may be used as a template variance structure for future times and a time series forecast (with suitable confidence bands) of the time index may then be used to generate a stochastic distribution of future rates, from which various statistics of interest may be calculated. The technique is very successful when dealing with mortality rates, in part because the 1-dimensional fit of the relational model explains almost all the variation in the process.

1 Lee-Carter: the main ideas

In my view the LC modeling enterprise involves the following steps. There are lots of variations on the details in how to implement each step, but the theme is unchanged across all LC-type models:

1.1 Data modelling and dimensionality reduction

In this step the goal is to find a "good" representation of the underlying historical series of central death rates $m(x,t)$ in terms of a simple model where age and time effects are separable.

$$f(m(x,t)) = a_x + k_t b_x \tag{1}$$

for some monotonic tranformation $f()$. With luck, this model captures most of the variation in the historical data and the age profiles a_x and b_x are *smooth* curves with a demographic meaning.

1.2 Modelling the time parameter k_t as a time series

The second step is to fix the age profiles, and look at k_t, the mortality index parameter. The simpler k_t turns out to be, the easier it will be to forecast or extrapolate to future values. In the case of developed countries, k_t turns out to be *nearly linear*. This is a remarkable feature: not only is the mortality dynamic nearly unidimensional with respect to age (the b_x "shape" of mortality change) but the "pace" of mortality change has been nearly constant over the historical period.

Although k_t is nearly linear, it does bounce around its trend line. Thus, the standard LC model for k_t is the *random walk with drift*

$$k_t = k_{t-1} + d + \delta_t \tag{2}$$

where d is the drift term or slope of the line and δ_t is a deviation or error term. The historical volatility of δ is seen in the LC model to be a best guide for its volatility in the future. This allows for a quantitative measurement of uncertainty in forecasts of k. In particular, k for s years into the future may be seen as

$$k_{T+s} = k_T + sd + \sqrt{s}\delta \tag{3}$$

where δ is a random variable with mean 0 and variance equal to the sample variance of the δ_t, and T is the final or launch year.

1.3 Forecasting mortality with confidence bounds

Using the stochastic forecast of k, the age profiles from (**??**) may be brought back into the picture and a set of stochastic m_{xt} generated. In turn, stochastic values of any life table function at a future time t may be derived.

2 OLS estimation of the Lee-Carter Model

This material is from Wilmoth(1993) "Compuation Methods for Fitting and Extrapolation the Lee-Carter Model of Mortality Change."

The LC model

The functional form of the Lee-Carter model may be expressed as:

$$f_{xt} = \ln(m_{xt}) = a_x + b_x k_t + \epsilon_{xt} \tag{4}$$

where

- m_{xt} is the observed age-specific death rate at age a and time t
- a_x, b_x, and k_t are model parameters and ϵ_{xt} is an error term

OLS fitting

- a_x is the average of $\ln(m_{xt})$ over time
- the scale on b_x and k_t is arbitrary, so constrain $\sum_x b_x^2 = 1$
- With the assumption that $E\epsilon_{xt} = 0$, the above two features mean that $\sum_t k_t = 0$
- fit by minimizing the sum of squared errors:

$$\sum_{xt} w_{xt}(f_{xt} - a_x - b_x k_t)^2 \tag{5}$$

  where w_{xt} is an observation weight.

- *Normal Equations*: the following normal equations may be derived from the first order condition, setting the derivatives of (**??**) to zero, and honoring the constraints $\sum_x b_x^2 = 1$ and $\sum_t k_t = 0$. They may be solved by solving the equations iteratively, repeating until convergence.

$$\hat{a}_x = \sum_t w_{xt}(f_{xt} - \hat{b}_x \hat{k}_t) / \sum_t w_{xt} \tag{6}$$

$$\hat{b}_x = \sum_t w_{xt} \hat{k}_t (f_{xt} - \hat{a}_x) / \sum_t w_{xt} \, \hat{k}_t^2 \tag{7}$$

$$\hat{k}_t = \sum_x w_{xt} \hat{b}_x (f_{xt} - \hat{a}_x) / \sum_x w_{xt} \, \hat{b}_x^2 \tag{8}$$

Alternatively, this may be solved using the Singular Value Decomposition of $\{f_{xt}\}$ or decomposing the variance-covariance matrix using principal components analysis (PCA). Other function minimization methods (e.g. Newton-Raphson) provide alternative ways to solve the set of equations.