Forecasting the Future: Using Lagged Data to Improve Coherent Mortality Forecasting

Introduction

Coherent forecasting takes the experience of two or more populations into account and ensures that the resulting forecasts for each population are 'non-divergent', which encompasses the conditions that they do not converge (and cross over) in the short term nor diverge in the long term (Li and Lee 2005). In recent applications, coherent forecasts have been based on the experience of a standard (or external reference population).

The use of an external standard in coherent mortality forecasting can improve the accuracy of the forecast, depending on the choice of standard. However, it is by no means clear how to choose *a priori* a standard that will be advantageous. Previously used standards include the mortality of the other sex in sex-specific forecasting, the mortality of the national population in subnational forecasting, and the mortality of a group of populations (Li and Lee 2005, Hyndman, Booth and Yasmeen 2013, Kjærgaard, Canudas-Romo, and Vaupel 2016, Bergeron-Boucher, Canudas-Romo, Oeppen and Vaupel 2017, Rabbi 2018).

Booth (forthcoming) found some evidence from sex-coherent forecasts that a low-mortality standard is advantageous, and hypothesised that in the context of mortality decline a low-mortality standard would serve as a good guide to the future. However, results using different low-mortality standards are mixed and it is clear that other features of the standard play a role. These features have not been identified, though it is likely that differing age patterns of decline are involved.

This paper further develops the use of a low-mortality standard by using forward lagged data of the population of interest as the standard. This approach is expected to remove any effects arising from differing age patterns of decline and hence produce more accurate forecasts.

Data

Data are obtained from the Human Mortality Database for the period 1950 to 2014. A total of 21 countries are included in the analysis, which is conducted separately for female and male mortality in recognition of the different mortality levels and patterns by sex. The data comprise annual age-specific central death rates by single years of age (with upper age group 95+) and corresponding populations exposed to the risk of death.

Method

The forecasts are made using the product-ratio coherent method with functional data models (Hyndman, Booth and Yasmeen 2013). The functional data model (Hyndman and Ullah 2007) is a generalisation of the Lee-Carter model:

$$\ln\big(m(x,t)\big) = a(x) + \textstyle\sum_j b_j(x)k_j(t) + e(x,t) + \sigma(x,t)\,\varepsilon(x,t) \qquad (1)$$

where $a(x)$ is the temporal average pattern of the logarithm of mortality by age and, for j = 1,...,J components, $b_j(x)$ is a 'basis function' and $k_j(t)$ is a time series coefficient. Broadly, the $k_j(t)$ represent annual rates of mortality decline averaged over age, while the $b_j(x)$ describe the age pattern of decline averaged over time. The parameters of the model are estimated after smoothing the data over age. Thus, the $a(x)$ and $b_j(x)$ are smooth functions of age. The pairs $(b_j(x), k_j(t))$ for j = 1,...,J are estimated using principal component decomposition. The error term $\sigma(x,t)\,\varepsilon(x,t)$ accounts for age-varying observational error; this is the difference between the observed rates and the smoothed rates. The error term $e(x,t)$ is modelling error, or the difference between the smoothed rates and the fitted rates from the model.

The product-ratio method for coherent forecasting uses the FDM in jointly forecasting mortality for two or more populations. The method is described here for sex-coherent forecasting. The product function is the geometric mean of sex-specific rates, $p(x,t) = \sqrt{m_F(x,t)m_M(x,t)}$, where F denotes female and M denotes male. The ratio function is the square root of the ratio of sex-specific rates, $r(x,t) = \sqrt{m_F(x,t)/m_M(x,t)}$. These functions are independently forecast using the FDM ($p(x,t)$ and $r(x,t)$ replacing $m(x,t)$ in Eq.1). Coherence is achieved by restricting the forecast of the ratio to converge very slowly to its temporal average; in other words, the forecast of each time coefficient becomes stationary.

The forecasts of the product and ratio functions are combined to produce forecast mortality rates. Forecast female mortality is the product of the forecasts of these two functions for future t:

$$\sqrt{\widehat{m_F(x,t)\,m_M(x,t)}}\,.\sqrt{\widehat{m_F(x,t)\,/\,m_M(x,t)}} = \sqrt{\widehat{m_F(x,t)^2}} = \widehat{m}_F(x,t) \qquad (2)$$

and forecast male mortality is their ratio:

$$\sqrt{\widehat{m_F(x,t)\,m_M(x,t)}}\,/\sqrt{\widehat{m_F(x,t)/m_M(x,t)}} = \sqrt{\widehat{m_M(x,t)^2}} = \widehat{m}_M(x,t). \qquad (3)$$

The product-ratio method makes use of the fact that the product and ratio will behave roughly independently of each other, as long as the two populations have approximately equal mortality variances (Hyndman et al 2013). The method is directly applicable to the mortality of any two populations for which the coherence of their future mortality is postulated.

In this paper, the two populations are sex-specific observed data for the period ending in year $T - L$, where T denotes the last year of observation and L denotes lag, and observed data for the same sex for the period ending in year T. The latter is regarded as the standard: it leads in time and in mortality decline. The product is thus the geometric mean referring to the year $T - L/2$. The ratio is the square root of the improvement over the lag or L-year period, in other words the improvement over L/2 years (assuming constant change over the lag). In order to obtain forecast rates referring to the same day within the year (30 June), lags of even numbers of years are appropriate. By using Eq. 2, forecast rates are obtained referring to years $T + 1, T + 2$, etc. This is true for any lag.

The evaluation involves comparing for each sex-country the forecast based on the lagged standard with the independent forecast and the sex-coherent forecast. The forecasts are evaluated for individual horizons 1 to 20, using a rolling fitting period in the calculation of average error so as to reduce the effect of fluctuation and abrupt change in annual mortality rates in relation to the fitting period and the forecast period. The evaluation of accuracy and bias is based on relative errors: respectively, the mean absolute relative error, $MARE$, and mean relative error, MRE, in age-specific mortality rates, averaged over age and fitting period. The use of relative errors gives equal weight across ages, regardless of the size of the rate.

The units of analysis for evaluation and comparison are $MARE(h, c)$ and $MRE(h, c)$. Horizon-specific mean accuracy and bias, $MARE(h)$ and $MRE(h)$, are averages over countries. Country-specific mean accuracy and bias, $MARE(c)$ and $MRE(c)$, are averages over horizons. Overall mean accuracy and bias, $MARE$ and MRE, are averages over countries and horizons. Standard deviations of the accuracy and bias measures, across horizons for each country and across countries for each horizon, are used to assess method robustness.

Results

Figure 1 shows mean accuracy averaged over age, fitting period and horizon for female mortality in the 21 countries by method. Overall means across countries show a reduction in MARE from 0.13 to 0.09 when using the lagged method with a lag of 10 years. Similar improvements are found for most measures. The results are summarised in Table 1. Results show that forecast accuracy and bias are generally improved by using a lagged standard with lag of 10 years, and that heterogeneity across countries is reduced.

Figure 1 Mean accuracy averaged over horizon for female mortality in the 21 countries by method

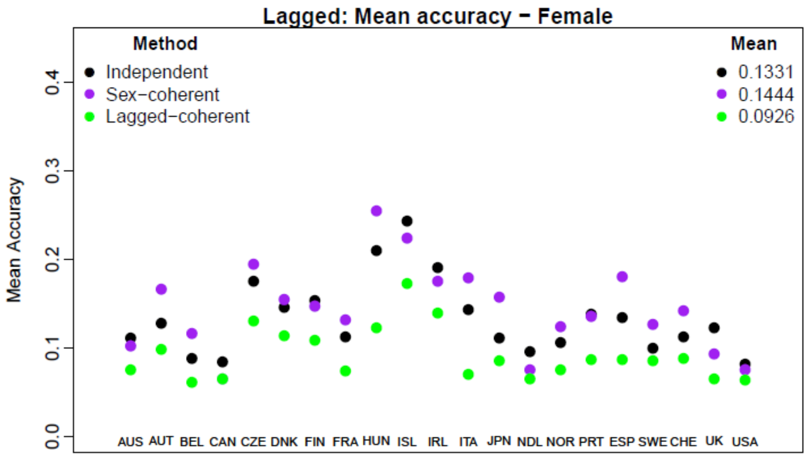

Table 1 Mean and variation in accuracy and bias relative to independent forecasts by method, female and male mortality

	Indep-endent	Sex-coherent	Lag-10
FEMALE			
Mean accuracy	1.00	1.08	0.70
SE(accuracy)	1.00	1.18	0.65
Mean bias	1.00	3.70	1.16
SE(bias)	1.00	1.27	0.69
MALE			
Mean accuracy	1.00	0.88	0.68
SE(accuracy)	1.00	0.77	0.69
Mean bias	1.00	0.58	0.64
SE(bias)	1.00	0.64	0.74

Next steps

Further work on this method is envisaged. It is possible (desirable) to repeat the forecast for several lags, averaging the result, thereby potentially increasing accuracy. The effect of lag length will also be examined.

References

Bergeron-Boucher, M.-P., Canudas-Romo, V., Oeppen, J., and Vaupel, J. W. (2017). Coherent forecasts of mortality with compositional data analysis. *Demographic Research*, 37:527–566.

Booth H. (Forthcoming 2020) Coherent mortality forecasting with standards, in S. Mazzuco and N. Keilman (eds) *New Perspectives on Population Forecasting: a presentation of new developments in the field of population forecasting*. Springer

Hyndman, R. J., Booth, H., and Yasmeen, F. (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography*, 50(1):261–283.

Kjærgaard, S., Canudas-Romo, V., and Vaupel, J. W. (2016). The importance of the reference populations for coherent mortality forecasting models. In *European Population Conference*, Mainz, Germany.

Li, N and Lee, R Li, N. and Lee, R. (2005). Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*, 42(3):575–594.

Rabbi, AMF (2018) Coherent mortality forecasting with Lee-Carter type models: Adjusting for smoothing and lifespan disparity. Unpublished PhD Thesis, University of Padua.