Insurance: Mathematics and Economics 77 (2017) 97–110

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

A class of random field memory models for mortality forecasting☆

P. Doukhan [a], D. Pommeret [b], J. Rynkiewicz [c], Y. Salhi [d,*]

[a] Université de Cergy-Pontoise, UMR 8088 Analyse, Géométrie et Modélisation, 95302 Cergy-Pontoise cedex, France
[b] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, 13288 Marseille cedex 9, France
[c] Équipe SAMM, EA 4543 Université Paris I Panthéon-Sorbonne, 90, rue de Tolbiac 75634 Paris cedex 13, France
[d] Univ Lyon, Université Lyon 1, ISFA, LSAF EA2429, 50 avenue Tony Garnier, F-69007 Lyon, France

ARTICLE INFO

Article history:
Received September 2016
Received in revised form July 2017
Accepted 27 August 2017
Available online 5 September 2017

MSC:
primary 60G70
secondary 60G10
60F99

Keywords:
Mortality rates
AR-ARCH random field
Estimation
QMLE
Inference

ABSTRACT

This article proposes a parsimonious alternative approach for modeling the stochastic dynamics of mortality rates. Instead of the commonly used factor-based decomposition framework, we consider modeling mortality improvements using a random field specification with a given causal structure. Such a class of models introduces dependencies among adjacent cohorts aiming at capturing, among others, the cohort effects and cross generations correlations. It also describes the conditional heteroskedasticity of mortality. The proposed model is a generalization of the now widely used AR-ARCH models for random processes. For such a class of models, we propose an estimation procedure for the parameters. Formally, we use the quasi-maximum likelihood estimator (QMLE) and show its statistical consistency and the asymptotic normality of the estimated parameters. The framework being general, we investigate and illustrate a simple variant, called the three-level memory model, in order to fully understand and assess the effectiveness of the approach for modeling mortality dynamics.

1. Introduction

The forecast of future mortality improvements poses a challenge not only for public retirement systems planning but also for the private life annuities business, due to the continuous longevity improvement. For public policy, as well as for the management of financial institutions, it is important to forecast future mortality rates in order to quantify the risk underlying their pension and annuities portfolios. To this end, a variety of models have been introduced in the literature during the last decades.

Most notably, there are the so-called factor-based models widely in use by practitioners, which know an increasing recognition from the actuarial community. These *traditional* mortality models rely on a factor-based decomposition of mortality surface. These factors are intended to capture the complex patterns of mortality evolution over time. Although these models are quite intuitive, their statistical properties are, however, not accurately understood. For instance, in their seminal and influential work, Lee and Carter (1992) have proposed a model that decomposes mortality surface into a latent trend, and two corresponding age-sensitive parameters, see also Brouhns et al. (2002). The other models that followed extend the idea underlying the Lee and Carter's (1992) model by adding a mixture of additional components which capture age, period and, in some cases, cohort effects. As noted by Mavros et al. (2017), "*the number and form of these types of effects is usually what distinguishes one model from another*". However, some recent works show their limits, e.g. Giacometti et al. (2012), Chai et al. (2013), Hunt and Villegas (2015) and Mavros et al. (2017) among others. In particular, one of the main drawbacks of these classical models relates to the assumption of the homoskedasticity of their residuals. In fact, the assumption of constant variance is always violated in practice as it is time varying, see e.g. Lee and Miller (2001) and Gao and Hu (2009). Furthermore, the mortality evolution is known to be related to the age of birth, see Willets (2004). This is generally referred to as the cohort effect and translates the persistent of some shocks on mortality among cohorts. It is observed when plotting the residuals of some models that rely on age and period factors as an apparent diagonal structure.

☆ The work is supported by the ANR Research Project LoLitA (ANR-13-BS01-0011). P. Doukhan's work has been developed within the MME-DII center of excellence (ANR-11-LABEX-0023-01). Y. Salhi's work is supported by the BNP Paribas Cardif Chair "Data Analytics & Models for Insurance". The views expressed herein are the authors' own and do not reflect those endorsed by BNP Paribas. We thank the editor and the anonymous referees for their careful reading, and their useful comments and suggestions which made our earlier presentation better. We would also like to thank Quentin Guibert for his useful comments.

* Corresponding author.
E-mail addresses: paul.doukhan@u-cergy.fr (P. Doukhan), denys.pommeret@univ-amu.fr (D. Pommeret), joseph.rynkiewicz@univ-paris1.fr (J. Rynkiewicz), yahia.salhi@univ-lyon1.fr (Y. Salhi).

http://dx.doi.org/10.1016/j.insmatheco.2017.08.010
0167-6687/

These observations point to a need for additional univariate cohort-dependent process in some countries. Such a phenomenon has lead to various extensions, in the literature, of the initial Lee–Carter model, e.g. Renshaw and Haberman (2006) or Cairns et al. (2009). The incorporation of the cohort-specific process, for instance, has been suggested to overcome the so-called *non-stationary* effect, which corresponds to the diagonal structure observed in the plotting of the age–period models' residuals. Even if this undesired remaining diagonal effect is, generally, accommodated, it is still unclear how such a cohort-effect can be interpreted and identified, see Hunt and Villegas (2015). This is even more appealing in view of some recent empirical findings. These praise the goodness-of-fit performance of age–period–cohort models specification but meanwhile shed light on their instable forecasting performance. Furthermore, these mainstream models are over-parameterized and have tendency to over-fit and thus produce less reliable forecasts.

It is of course very important to tackle these limitations when considering a new modeling approach, but it is also essential to take into account the dependence structure between adjacent cohorts. Indeed, some recent works, and even common intuition, point out the importance of cross-cohorts correlation, see e.g. Loisel and Serant (2007) and Jevtić et al. (2013). In their empirical work, Loisel and Serant (2007) show that correlation among close generations is higher enough to be omitted. The same conclusions were drawn in the very recent work of Mavros et al. (2017).

In this paper, in contrast to this univariate factor-based framework, we approach the problem of modeling mortality rates by considering the whole surface of mortality improvements as a sole random field without any further assumption on the particular dependence structure neither the factors driving its evolution. Thus, unlike mainstream approach, our modeling framework is intended to accommodate cross-cohorts dependence as well as conditional heteroskedasticity. The starting point of our approach is a formulation of the mortality random field in the sense of Doukhan and Truquet (2007) with a given causal structure. Such a class of models introduces dependencies among adjacent cohorts aiming at capturing, among others, the cohort effects and cross-generations correlations. It also takes into account the conditional heteroskedasticity of mortality. The proposed model is a generalization of the now widely used AR-ARCH models for random processes. More formally, the conditional mean and variance of mortality rates are respectively described by linear combinations of the observed rates and their squared values on a given neighborhood. In Section 2, we fully describe the model and give some intuitions on its construction. The specification of the causality structure is discussed and some first results on the stability as well as the identification of the model are introduced. For such a class of models, we also propose a robust estimation procedure for the parameters.

The rest of the paper is organized as follows. In Section 3, we use the quasi-maximum likelihood estimator (QMLE) to estimate the parameters. Its statistical consistency and asymptotic normality are shown. The framework being general, we investigate and illustrate a simple variant, called the three-level memory model, in order to fully understand and assess the effectiveness of the approach for modeling mortality dynamics. This three-level memory level incorporates the correlations with the immediate cohorts and it is intended to capture the cohort effect in a natural manner. In Section 4, the model is applied to the populations of US, France and England & Wales, and is compared to the benchmark models of Lee and Carter (1992) and Cairns et al. (2006) two-factor models.

2. Random fields memory models

2.1. From classical mortality models to a random field memory formulation

Denote by $m_{(a,t)}$ the crude death rate at age a and date t. Time is assumed to be measured in years, so that calendar year t has the meaning of the time interval $[t, t+1)$. For expository purpose and since we will be working with only a subset of historical data, we will henceforth re-index the observable ages by $a = 0, 1, \ldots, I-1$ and the observable dates by $t = 0, 1, \ldots, J-1$; where I and J are, respectively the number of ages and years. Here, we introduce two benchmark models for mortality dynamics in order to motivate the development of the random field model discussed later on this section. We limit ourselves to these models for simplicity and other modeling frameworks are briefly discussed.

Classical mortality models. In their seminal paper, Lee and Carter (1992) postulated that the (log) mortality rates at different ages are captured by a common factor, and an age-specific coefficient with respect to this common trend. More precisely, we have for any a and t

$$\log m_{(a,t)} = \alpha_a + \beta_a \kappa_t + \epsilon_{(a,t)}, \text{ with } \epsilon_{(a,t)} \sim \mathcal{N}(0, \sigma) \tag{1}$$

where α_a is the time average level of $\log m_{(a,t)}$ at age a, κ_t is the common factor also known as the period mortality effect and β_a is the age-specific sensitivity coefficient with respect to κ_t. Another interesting model was suggested by Cairns et al. (2006) and assumes that the one-year death mortality rates dynamics are given by the following modeling form:

$$\operatorname{logit} q_{(a,t)} = \kappa_t^{(1)} + \kappa_t^{(2)}(a - \bar{a}) + \epsilon_{(a,t)}, \text{ with } \epsilon_{(a,t)} \sim \mathcal{N}(0, \sigma) \tag{2}$$

where $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ are two time varying stochastic period factors and $\bar{a}$ is the mean of the ages in the data. The innovation $\epsilon_{(a,t)}$ is assumed to be drawn from an i.i.d. zero-mean Gaussian random variable with constant variance σ^2. Such models describe the principal mortality dynamics in the sense that it includes the age related basis component and all of the non-stationary stochastic part of the mortality surface. The time-dependent parameters in both models are generally modeled using a simple ARIMA(0,1,0) model. On the other hand, recent works on mortality demonstrated the existence of the so-called cohort effect which makes the mortality depend not only on the age and calendar year but also on the year of birth. Over all, this advocates the inclusion of an additional factor γ_{t-a} being dependent on the year of birth $t-a$, see Renshaw and Haberman (2006), Cairns et al. (2009) and Hunt and Villegas (2015).

One of the drawbacks of these classical models relates to the assumption of homoskedastic error terms $\epsilon_{(a,t)}$. In fact, in practice, the assumption of constant variance is always violated: the observed logarithm of central death rates is much more variable and the volatility is time varying, see e.g. Lee and Miller (2001) and Gao and Hu (2009). Furthermore, the mortality evolution is known to be related to the age of birth, see Willets (2004). This is referred to as the cohort effect and translates the persistent of some shocks on mortality among cohorts. This is generally observed when plotting the residuals $\epsilon_{(a,t)}$ of models (1) and (2) as an apparent diagonal structure which requires additional univariate cohort processes in some countries. As noted above, this phenomenon has led to various extensions of the initial Lee–Carter model by introducing factors γ_{t-a} dependent on the year of birth, e.g. Renshaw and Haberman (2006) or Cairns et al. (2009) and the reference therein. However, the inclusion of additional univariate processes enhances the goodness-of-fit of the model but *over-fit* the data and thus produces less reliable forecasts, see Hunt and Villegas (2015).