J. Math. Study
doi: 10.4208/jms.v57n4.24.06

Vol. **57**, No. 4, pp. 486-498
December 2024

Fitting Heavy-Tailed Distributions to Mortality Indexes for Longevity Risk Forecasts

Longyu Chen[1], Tsz Chai Fung[2], Yinhuan Li[3,*]and Liang Peng[2]

[1] *School of Mathematics, Jilin University, Jilin 130012, China;*
[2] *Maurice R. Greenberg School of Risk Science, Georgia State University, Georgia 30303, USA;*
[3] *School of Statistics, East China Normal University, Shanghai 200241, China.*

Received March 23, 2024; Accepted May 6, 2024;
Published online December 13, 2024.

Abstract. Modeling and predicting mortality rates are crucial for managing and mitigating longevity risk in pension funds. To address the impacts of extreme mortality events in forecasting, researchers suggest directly fitting a heavy-tailed distribution to the residuals in modeling mortality indexes. Since the true mortality indexes are unobserved, this fitting relies on the estimated mortality indexes containing measurement errors, leading to estimation biases in standard inferences within the actuarial literature. In this paper, the empirical characteristic function (ECF) technique is employed to fit heavy-tailed distributions to the time series residuals of mortality indexes and normal distributions to the measurement errors. Through a simulation study, we empirically validate the consistency of our proposed method and demonstrate the importance and challenges associated with making inferences in the presence of measurement errors. Upon analyzing publicly available mortality datasets, we observe instances where mortality indexes may follow highly heavy-tailed distributions, even exhibiting an infinite mean. This complexity adds a layer of difficulty to the statistical inference for mortality models.

AMS subject classifications: 62M10, 62P05

Key words: Characteristic function, Lee-Carter model, mortality rates, unit root model.

---

1 Introduction

The increasing life expectancy has placed significant pressure on managing longevity risk for pension funds. Modeling mortality is crucial for comprehending mortality patterns, predicting mortality risk, hedging longevity risk, and pricing life insurance products like

*Corresponding author. *Email addresses:* `chenly19@mails.jlu.edu.cn` (Chen L), `tfung@gsu.edu` (Fung T C), `52204404004@stu.ecnu.edu.cn` (Li Y), `lpeng@gsu.edu` (Peng L)

annuities. One benchmark mortality model is the well-known Lee-Carter model introduced by [12]. This model captures central death rates in two steps: first, it fits a linear model to the logarithms of death rates with an unobserved mortality index k_t given by

$$\log m(x,t) = \alpha_x + \beta_x k_t + \varepsilon_{x,t}, \qquad \sum_{x=1}^{M} \beta_x = 1, \qquad \sum_{t=1}^{T} k_t = 0, \tag{1.1}$$

and second, it

$$\text{fits a time series model to } \{k_t\} \text{ with residuals } \{e_t\}, \tag{1.2}$$

where $x = 1,\ldots,M$ and $t = 1,\ldots,T$ represent age groups and (calendar) time, respectively. Here, k_t serves as the time-varying mortality index that is uniform across all age groups, encapsulating the overall mortality trends, and $\varepsilon_{x,t}$ represents the age-specific error term with a zero mean.

The classical Lee-Carter model has been extensively explored and applied in actuarial science, as evidenced by works such as [2,5,6,9,16,18,21,22]. The two-step statistical inference approach proposed by [12] involves initially estimating α_x, β_x, and k_t through singular value decomposition (SVD) applied to (1.1). Subsequently, an ARIMA(p,d,q) model is fitted to the estimated mortality indexes from the first step. Notably, many research works adopt a unit root AR(1) model as a special case of the general ARIMA(p,d,q) model when investigating longevity risk within the Lee-Carter framework. This is evident in studies such as [1,14,15,17].

When modeling and forecasting mortality rates, the influence of extreme mortality events, such as the influenza epidemic in 1918, World War I and II, and the recent COVID-19 pandemic, cannot be overlooked. Consequently, the incorporated mortality model must account for the potential impacts of such extreme scenarios. This is crucial for generating accurate forecasts of future mortality indexes and, consequently, assessing longevity risk. A straightforward method to address this is to fit parametric heavy-tailed distributions to the residuals e_t in modeling k_t. However, the literature on heavy-tailed modeling for mortality indexes is relatively sparse. [4] integrates the unit root AR(1) model for k_t with a Generalized Pareto Distribution (GPD) for the residuals by modeling mortality improvement under a threshold with a normal distribution and beyond the threshold with the GPD. Similar approaches, integrating non-Gaussian innovations or Extreme Value Theory (EVT) into the Lee-Carter model, have been explored by [23] and [10]. Another approach is to model the unobserved mortality indexes by a jump model, see [3].

To fit these mortality models, researchers often estimate the unobserved mortality indexes k_t's by applying the SVD method or others to (1.1), and then use the estimated mortality indexes to fit a time series model. Regrettably, this two-step statistical inference is problematic as noticed by [13]. Intuitively, the estimated mortality index $\tilde{k}_t$ via the SVD method or others may be decomposed into

$$\tilde{k}_t = k_t + \tilde{\varepsilon}_t \text{ with } \tilde{\varepsilon}_t \text{ depending on } \{\varepsilon_{x,t}\}_{x=1}^{M}. \tag{1.3}$$