FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 7 No. 3, June (Special Issue), 2023, pp 293 - 296
DOI: https://doi.org/10.33003/fjs-2023-0703-1874

ASSESSING THE SUITABILITY OF THE LEE-CARTER MODEL IN MODELLING MORTALITY DATA WITH VARYING b_x PARAMETER

[*1]Shelleng A. U., [2]Dikko H. G., [2]Garba J., [3]Alhaji B. B.

[1]Department of Mathematics, Gombe State University, Gombe
[2]Department of Statistics, Ahmadu Bello University Zaria, Kaduna
[3]Department of Mathematics, Nigerian Defence Academy, Kaduna

*Corresponding authors' email: Umssie@gmail.com

ABSTRACT

The Lee-Carter model, developed by Lee and Carter in 1992 is one of the most influential model among others that is used for mortality projection. Although the model's performance has so far been examined in a variety of situations, its effectiveness in modeling mortality data with varying speeds of change in mortality b_x across ages and the ability to detect trends in mortality index k_t more precisely over time has not been studied. The method traditionally used in the Lee-Carter model shows obvious drawbacks in describing the shape of future mortality. The effectiveness of the Lee-Carter model to model mortality data with varying speeds of change in mortality b_x was investigated. The model was applied to mortality data with both constant and varying speeds of change for female sub population. Results show that the Lee-Carter model is not good for mortality data with varying speeds of change.

Keywords: Lee-Carter model, ARIMA, Singular value Decomposition and Mortality Rate

INTRODUCTION

Understanding the assumptions of future mortality and longevity risk is crucial for people in actuarial science and the government in general. The welfare and productive life expectancies of millions of people are increased by longevity breakthroughs, but these same advancements also increase the cost of pension systems, endangering the long-term creditworthiness of financial institutions owing to upsurges in unexpected future obligations. Additionally, public health expenses are impacted if unhealthy life expectancy is prolonged as a result of increases in mortality rates at later ages. Both academics and professionals have realized the significance and necessity of accurately studying future mortality in order to deal with these rapid changes and avoid undesirable consequences. There has been a lot of work put into researching and creating more accurate mathematical models and methods for predicting mortality.

The need to forecast into the future and have a good understanding of what future mortality rates and life expectancy will look like in the future should be a major importance to researchers. Numerous methods have been developed recently to project mortality using stochastic models (McNown and Rogers, 1989; Lee and Carter, 1992; Alho, 1990; Currie *et al.*, 2004; Cairns *et al*., 2006; Renshaw and Haberman, 2006). The Lee-Carter (LC) model, created by Lee and Carter in 1992, is one of the most often used models for projecting mortality rates among the several models that have been used to forecast mortality. The LC model has garnered a lot of interest and has come to be regarded as the benchmark for mortality modeling and projection. Even though the model has received a lot of criticism, it continues to dominate mortality forecasts.

The LC model uses linear extrapolations of the age-specific death rates to project mortality. In particular, a logged matrix of age-specific mortality rates over time, a_x and $b_x$describes the mortality profile and the constant improvement of mortality at age x respectively, and also a time varying index k_t that captures the level of mortality generally. The projected of k_t is used to generate forecasts using conventional time series approaches. One major weakness of the LC model is the assumption of a fixed age-component over time, i.e., the change of mortality over all ages is constant. The mortality of the industrialized countries of the world has considerably shown age-time interaction. Therefore, it is appropriate to note that these two dominant trends have impacted different age groups.

The LC model continues to play a leading role in mortality forecasts, despite the fact that there have been many models employed for mortality forecasting throughout the years. The assumption of a fixed age component over time, however, is a significant issue with LC model (Lee and Miller 2001). This assumption of constant mortality variability over time seems to be unlikely and unworkable. Of course, we believe that the mortality trends seen in various countries over time show a significant age-time interaction, with mortality affecting different age groups at different times. Horiuchi and Wilmoth (1995) discovered that mortality reduces more quickly as people get older than at younger ages in various countries around the world. In their research, Lee and Carter (1992) did not consider this possibility. This research suggests that a fixed change in mortality may not adequately describe the mortalities of developing nations like Nigeria.

MATERIALS AND METHODS

The World Population Prospect (WPP) database of the United Nations was used to collect the data for this study. From 2009 to 2020, age-specific mortality rates data for Nigerian females were utilized for both overall mortality and piecewise mortality (Adults).

Lee-Carter model

Lee and Carter (1992) proposed a method for projecting future mortality rates. The main components of the Lee-Carter approach are the base model and the Autoregressive Integrated Moving Average (ARIMA) model. In the first stage, a_x, b_x and k_t are estimated using the actual mortality surface. In the second stage, the Box and Jenkins (1970) method is used to model and extrapolate the fitted values of k_t using an autoregressive integrated moving average process.

$ln(m_{xt}) = a_x + b_x k_t + \mathcal{E}_{xt}$ (1)

where m_{xt} is given as the matrix of the observed age-specific death rate at age *x* during year t.

a_x is the average shape of mortality by age.

b_x is the relative speed of change or variability in mortality by age as k_t changes.

k_t is the time trend for the general mortality. It is also be referred to as the mortality index.

ε_{xt} is the error term or residual at age x and time t.

Singular value Decomposition

Singular Value Decomposition (SVD) was introduced in 1991 by Bell and Monsell. It is a process for decomposing a matrix into each of its several constituent matrices, revealing many of the interesting and valuable properties of the original matrix. Its foundation is a linear algebraic theorem that states that a matrix Z can be reduced to the product of three matrices: an orthogonal matrix B, a diagonal matrix Σ, and the transpose of an orthogonal matrix K. The matrix should ideally be divided into a set of factors that are optimal according to some criterion. Matrix Z can be uniquely decomposed as:

$Z = B\Sigma K^T$ (2)

The matrices B and K are both defined to be orthogonal matrices. Matrix Σ is a diagonal matrix. The elements along the diagonal of Σ are known as the singular values of the matrix Z. The columns of B are known as the left-singular vectors. The right-singular vectors are the columns of K.

Since there are no provided regressors on the right side of the estimate, the LC model cannot be fitted using conventional regression techniques. The estimators of a_x, β_x and k_t will be denoted by $\hat{a}_x, \hat{\beta}_x$ and $\hat{k}_t$ respectively. Firstly, lets define the $\hat{a}_x$ estimate as the averaging $\ln(\hat{m}_{x,t})$ over a period of time t

$\hat{a}_x \frac{1}{n} = \sum ln\, m_{xt}$ (3)

The parameters β_x and k_t are calculated by applying Singular Value Decomposition to the matrix Z. where;

$Z_{xt} = ln\, m_{xt} - \hat{a}_x$ (4)

That is;

$SVD(Z_{xt}) = B\Sigma K^T$ (5)

Where B represents the age component, Σ represents the singular values and K represents the time component. Therefore after decomposition,

$\text{SVD}\,(Z_{x,t}) = \lambda_1 B_{x,1} K_{t,1} + \lambda_2 B_{x,2} K_{t,2} + \cdots + \lambda_i B_{x,i} K_{t,i}$
$= \sum_{i=1}^{r} \lambda_i B_{x,i} K_{t,i}$ (6)

Where r = Rank $|Z|$ and $\lambda_i (i = (1,2, \ldots, r)$ are the ordered singular values with $B_{x,i}$ and $K_{t,i}$ as the corresponding left and right singular vectors. These give the estimates of b_x *and* $k_t$as follows; $\hat{b}_x = U_{x,1}$ and $\hat{k}_t = V_{t,1}$

RESULTS AND DISCUSSION

The Lee-Carter model was applied to female mortality data to examine the impact on various age groups.

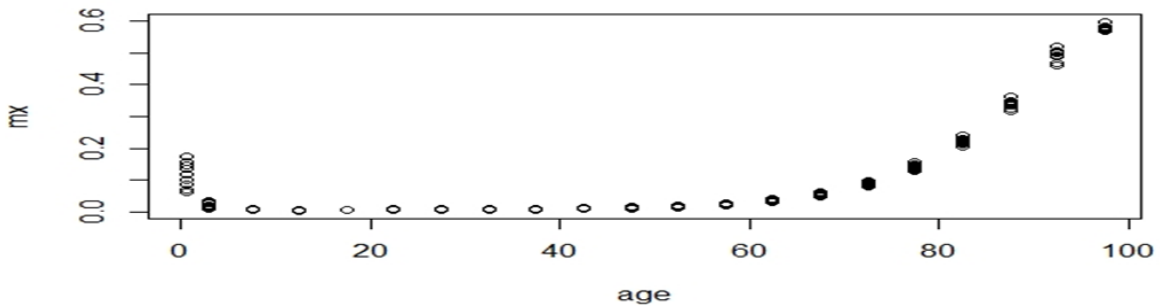

Figure 1: Central death rates m_x using the LC model

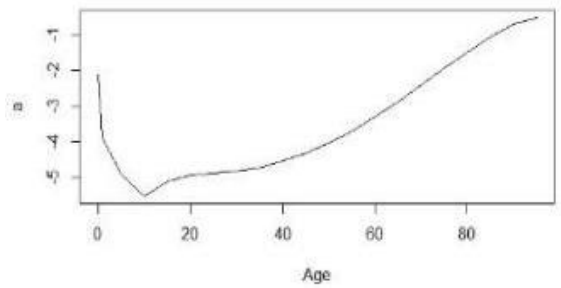

Figure 2: General age profile

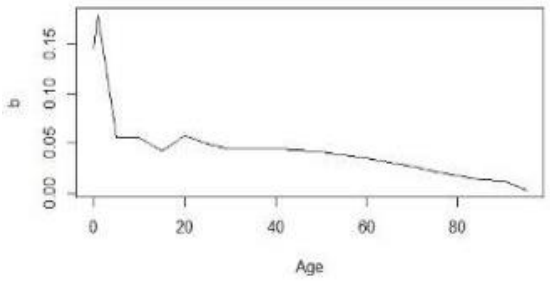

Figure 3: Pace of Mortality change

As seen in figures 1, 2 and 3, there was no regular pattern emerging from the plots, but rather we had non-linear trends. It is observed that the variability of mortality is not the same across all ages. Also, the contribution of some particular age groups, i.e. infant, child, and mortality at old age, to the overall mortality of the population is different. The high incidence and prevalence of childhood diseases and deaths at old age could have been majorly responsible for the changes in the shape of the parameter. Thus, as against what is obtainable in the original Lee-Carter model where the overall mortality is studied from age 0, the study will also consider piecewise or segmented mortality (adult mortality) to check the assumption of constant variability in mortality change. Therefore, our proposed model is given as:

$ln(m_{xt}) = a_x^i + b_x^i k_t^i + \varepsilon_{xt}^i$ (7)

$m_{xt} = exp^{m_{xt}^i}$ (8)

$i =$ infant, child, adult or old age mortality

LC Model for overall mortality profile with varying $b_x$parameter

Here the LC model is given as $m_{xt} = exp(a_x + b_x k_t + \varepsilon_{xt})$ where $x = 0 - 1{,}1 - 4{,}5 - 9, \ldots 100+$ and $t = 2009{,}2010, \ldots 2020$. The entire mortality profile using the original LC model by applying the following steps

i. Obtain the logarithm of mortality rates that is $ln(m_{xt})$
ii. Obtain the parameter $\hat{a}_x$which is a column vector, it is calculated as average over time of the logarithm of mortality rates $\hat{a} = \frac{1}{n}\sum_{t=1}^{n} ln(m_{xt})$
iii. Create a matrixz_{xt} for estimating $b_x$and k_t where $z_{xt} = (ln\, m_{xt}) - \hat{a}_x$
iv. Decompose the matrix z_{xt} into a product of matrices using the singular value decomposition method $SVD(z_{xt}) = B\Sigma K^T$

After applying the techniques stated in above, we obtained the following result:

$$\hat{a}_x = \begin{pmatrix} -2.12737 \\ -3.87952 \\ -4.8853 \\ -5.5063 \\ -5.09848 \\ -4.92957 \\ -4.8682 \\ -4.82472 \\ -4.71039 \\ -4.50945 \\ -4.31112 \\ -4.02745 \\ -3.70947 \\ -3.30332 \\ -2.89672 \\ -2.42628 \\ -1.9513 \\ -1.4972 \\ -1.07606 \\ -0.70793 \\ -0.54755 \end{pmatrix} \quad \hat{b}_x = \begin{pmatrix} 0.138443582 \\ 0.174867719 \\ 0.06519156 \\ 0.064015942 \\ 0.052109864 \\ 0.048287934 \\ 0.045804646 \\ 0.044942453 \\ 0.043885289 \\ 0.041910333 \\ 0.039741495 \\ 0.03812534 \\ 0.03611515 \\ 0.034482034 \\ 0.031256629 \\ 0.0274979 \\ 0.023421014 \\ 0.018904223 \\ 0.014094418 \\ 0.010414207 \\ 0.006488269 \end{pmatrix} \quad \hat{k}_t = \begin{pmatrix} 3.6266293 \\ 2.631595 \\ 1.5241754 \\ 0.6414605 \\ 0.6821964 \\ 0.7184163 \\ 0.9613649 \\ 0.7036283 \\ -0.8322227 \\ -2.0657114 \\ -3.5070368 \\ -5.0844953 \end{pmatrix}$$

From the decomposition of matrix z_{xt} we got the parameter $\hat{k}_t$ which is the mortality index, it is a vector that captures the overall mortality change overtime. A standard time series model selection procedure was carried out to identify the best ARIMA model for Nigerian mortality data using AIC, BIC and some error measures.

Table 1: AIC's, BIC's and Error measures for suggested models using female data

Model	AIC	BIC	RMSE	MAE	MAPE
ARIMA (0,1,0)	33.77257	34.17046	0.9819293	0.7795467	53.12297
ARIMA (1,1,1)	24.71286	25.90654	0.5160594	0.3924597	37.91306
ARIMA (1,1,0)	**23.01995**	**23.81574**	**0.5130491**	**0.3651628**	**36.37704**

It can be seen that from Table 1, ARIMA (1,1,0) seems to be the best ARIMA model for female Nigeria mortality data as against a random walk (0,1,0) that was used by Lee and Carter in 1992 and also widely used by other researchers. Thus, ARIMA (1,1,0) was used to model Nigeria's female mortality data for both the overall mortality profile and piecewise with specific age intervals.

Table 2: Forecasted mortality index $\hat{k}_t$ for overall mortality using ARIMA model

Year	2021	2022	2023	2024	2025	2026	2027
$\hat{k}_t$	-6.480502	-7.715928	-8.809246	-9.776801	-10.633061	-11.39082	-12.061427

Table 2 above shows the forecasted values of $\hat{k}_t$ for overall female mortality profile. The LC model is then updated with these predicted values to project future age-specific mortality. Next is to compare the performance of LC model on mortality data with varying $b_x$parameter for both overall mortality data and adulthood mortality.

LC Model for adult mortality data with constant $b_x$parameter

Here the LC model is given as $m_{xt} = exp(a_x + b_x k_t + \varepsilon_{xt})$ where $x = 15-19, 20-24, 25-29, \ldots 55-59$ and $t = 2009, 2010, \ldots 2020$. The model parameters $\hat{a}_x, \hat{b}_x$ and $\hat{k}_t$ were estimated using the SVD.

$$\hat{a}_x = \begin{pmatrix} -5.098475 \\ -4.929566 \\ -4.868200 \\ -4.824719 \\ -4.710387 \\ -4.509445 \\ -4.311115 \\ -4.027447 \\ -3.709472 \end{pmatrix} \quad \hat{b}_x = \begin{pmatrix} 0.13263363 \\ 0.12408778 \\ 0.11800242 \\ 0.11591519 \\ 0.11291646 \\ 0.10708573 \\ 0.10084541 \\ 0.09680519 \\ 0.09170820 \end{pmatrix} \quad \hat{k}_t = \begin{pmatrix} 1.7429969 \\ 1.0784067 \\ 0.3366126 \\ -0.2415136 \\ -0.2147332 \\ -0.1913032 \\ 0.3662920 \\ 0.8638399 \\ 0.1940775 \\ -0.4626098 \\ -1.2300380 \\ -2.2420278 \end{pmatrix}$$

Table 3: Forecasted mortality index $\hat{k}_t$ for Adulthood mortality using ARIMA model

Year	2021	2022	2023	2024	2025	2026	2027
$\hat{k}_t$	-3.016034	-3.608023	-4.06080	-4.407095	-4.671957	-4.87453	-5.02947

Table 3 shows the forecasted values of $\hat{k}_t$ for adulthood mortality. The LC model is then updated with these predicted values to project future age-specific mortality.

Table 4: Comparison between m_x (Overall mortality profile) and m_x (Adulthood mortality)

Mortality	MSE	RMSE	MAE	MAPE
Overall mortality	0.0003216921	0.0179357	0.009359309	0.1920557
Adulthood mortality	$2.696898e^{-09}$	$5.193167e^{-05}$	$\mathbf{3.06832e^{-05}}$	0.003220681

Results in Table 4. Shows that piecewise mortality (adulthood mortality) has minimum errors as compared to overall mortality.

CONCLUSION

Though the performance of the LC model in modelling the entire mortality profile has been studied in some literature (Li and Chan, 2007; Nan et al. 2015; Taruvinga et al. 2017), there is no concrete research in terms of piecewise mortality modelling for different ages. For this reason, in this study, the LC model was utilized within the perspective of overall mortality modelling and piecewise modelling. We take into consideration the demographic data for Nigeria. Results show noticiable defference between the two approaches to modelling moratlity data with varying $b_x$and moratlity data with constant b_x. As seen in Table 4, the adulthood mortality modelling approach gives minimum errors with regards to the forecast of the mortality rate of age 15 to 59 for female population. Mortality data with a constant $b_x$gives a more accurate forecast in terms of mortality rates. This supports our suggestion that for some mortality data, the whole age span has a varying speed of change in mortality and thus may not be modelled well by the original LC model.

REFERENCES

Alho, J.M. (1990). Stochastic methods in population forecasting. *International Journal of Forecasting* 6(4), 521–530.

Bell, W. and Monsell, B. (1991). Using principal components in time series modeling and forecasting of age-specific mortality rates. *Alexandria Virginia American Statistical Association*.

Box, G.E. and Jenkins, G.M. (1970). *Time series analysis: Forecasting and control.* San Francisco: Holden-Day.

Cairns, A.J., Blake, D. and Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance 73*(4), 687–718.

Currie, I. D., Durban, M. and Eilers, P.H (2004). Smoothing and forecasting mortality rates. *Statistical Modelling 4* (4), 279–298.

Horiuchi, S. and John, R.W. (1995) "The aging of mortality decline" Presented at the Annual Meeting of the Population Association of America, San Francisco, April 6-8, 1995.

Lee, R. and Carter, L. (1992). Modeling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association* 87: 65-71.

Lee, R. and Miller, T. (2001). Evaluating the performance of the Lee-Carter method for Forecasting mortality *Demography*, Vol. 38, No. 4 (Nov., 2001), pp. 537-549.

McNown, R. and Rogers, A. (1989). Forecasting mortality: A parameterized time series Approach. *Demography 26*(4), 645–660.

Renshaw, A.E., and Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for Mortality reduction factors. *Insurance: Mathematics and Economics* 38(3): 556-570