Original Research Article

Stochastic Mortality in Kenya

Abstract

This research work seeks to analysis the mortality trend Kenya has experienced from 1950 to date and the expected mortality characteristics. Age specific period life table functions are utilized to analyze life expectancy evolution, age at death distribution, and age survival rates. Life expectancy at any given age has increased. Kenya has shifted from a high mortality population to an intermediate state, characterized by reduced infant mortality. The modal age at death has slightly reduced to 73 years. The rate of survival has also increased. Gompertz model is estimated to to evaluated the force of mortality of adults thus provide evidence of 6% annual increment starting from 35 person per ten thousand at the age of 25 years. Gini mortality coefficient and Lorenz curve are used to study the life disparity experienced. The inequality in lifespan is moderately high with slight improvement. Lastly, Lee-Carter model is estimated which helps in forecasting life expectancy. The model turn out to be statistically significant at 5% significant level. The life expectancy is expected to increase with time. The projected life expectancy at birth is 65.6 and 70.5 years for the year 2030 and 2071 respectively.

Keywords and phrases

Life expectancy,mortality rate,life table, age specific mortality rate

1 Introduction

Mortality is the number of deaths occurring during in a population during a given time period. There are several mathematical ways of analyzing mortality. The basic way is the crude death rate, defined as the proportion of deaths to the population at a given time period. A alternative way to see the death rate is by considering the death rate at exact age. The proportion of individuals death at an exact given age to its respective population at same exact given age is called age specific mortality rate. Life expectancy is the estimated number of year a person at exact given age is expected to live. Life expectancy assumes that the age specific death rates in a particular years applies through out an individual given aged life time, thus its a hypothetical measure. Globally in the last century, the has been a reduction in mortality and increase in life expectancy. These improvement have been attributed to awareness of health behavior and modernization of health infrastructure (Roser et al. 2013). Global life expectancy at birth in 2021 is 71 years with the least developed countries lagging behind by 7 years (UN WPP 2022). Covid-19 pandemic period between late 2020 and 2021 resulted to slowed life expectancy improvement in the world. Life expectancy at birth in the sub-Saharan Africa is 59.7 for both sexes. Female lead with 61.6 years as compared to males at 57.8 years.

Age specific period life table function are tools used to analyse mortality. The shape of age-at-distribution derived from the life table death density function, is used to describe variablity of age at death, shift in modal age at death and any shift in mortality (Canudas-Romo 2008,Canudas-Romo 2010,Ouellette et al.,Horiuchi et al. 2013, Basellini et al. 2019, Shang and Haberman 2020). Survival curves dimensions are used in mortality analysis to determine highest normal life duration exceeds modal age, death around the modal age and the proportion of survivors in population (Cheung et al. 2005, Ebeling et al. 2018). Gini coefficient and Lorenz curve are econometric toolkit for inequality measure. However, recently they have been adapted to determine ineqaulity of life. Lorenz curve shows the spread of the disparity from the equality line while Gini coefficient quantify the area between line of equality and Lorenz curve. Shkolnikov (2003) developed a framework of applying the two measure to evaluate life inequality. This techniques was been applied by Vaupel et al. (2011) ,Giorgi et al. 2017 and Zafeiris 2023 in their disparity analysis. In mortality modeling and forecasting, the model developed by Lee and Carter (1992) has been widely used. Lee-Carter model is time series model to forecast long-run age specific mortality in US. Age specific mortality is expressed as function unobserved time specific intensity index and an additional parameter dependent on age. Several other studies have been conducted using this model in different countries such as; Chile (Lee and Rofman 1994), G7 countries (Tuljapurkar et al. 2000), Australia (Booth et al. 2002, Carter and Prskawetz 2001, and Booth and Tickle 2003), Sweden (Lundström and Qvist 2004), Spain (Debón et al. 2008), India (Chavhan and Shinde 2016), Malaysia (Ibrahim et al. 2021) and Bangladesh (Fazle and Khan 2022)

The aim of this paper is to analysis the trend and model mortality in Kenya. This is achieved by analyzing the life table function to depict underlying trend in mortality characteristic and estimation of future life expectancy.

2 Methodology

2.1 Data

The source of data employed in this study is the UN (United Nation) WPP22 (world population prospect year 2022) [23]. The dataset consist of single year population and deaths for both male and female from 1950 to 2021. According to Keilman (1998) research o the accuracy of UN world population projection using 12 set of population projection between 1950 and 1980, he observed a clear tendency over time towards accuracy improvements.

Crude mortality rate trend in Kenya is shown figure 1. Its evident that mortality rate has reduced over time. There was however a break in the decrease trend in the 1990s. Age specific mortality (ASMR) are analyzed to show clear variation in mortality across age and year as shown in figure 2. Infant mortality has sharply declined over time. From 6 years to 63 year, there has been less variability in ASMR over time. After 63 years, ASMR tends to increase though there has been a decrease in value over time. In the neighborhood of 100 years there is a lot of volatility over time.

Figure 1: crude mortality rate plot

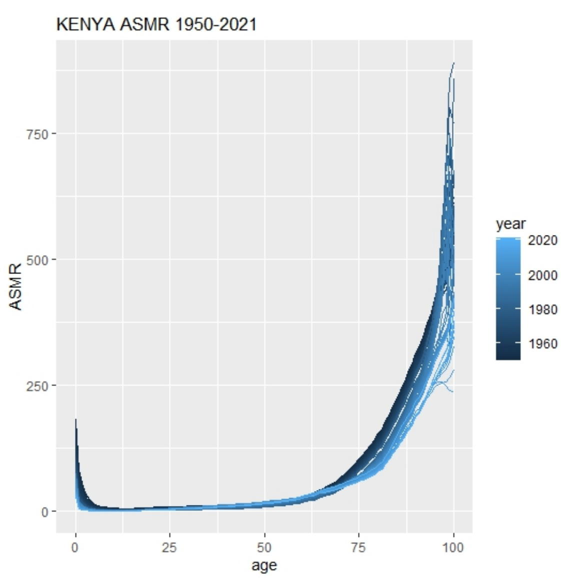


Figure 2: ASMR plot

2.2 Model

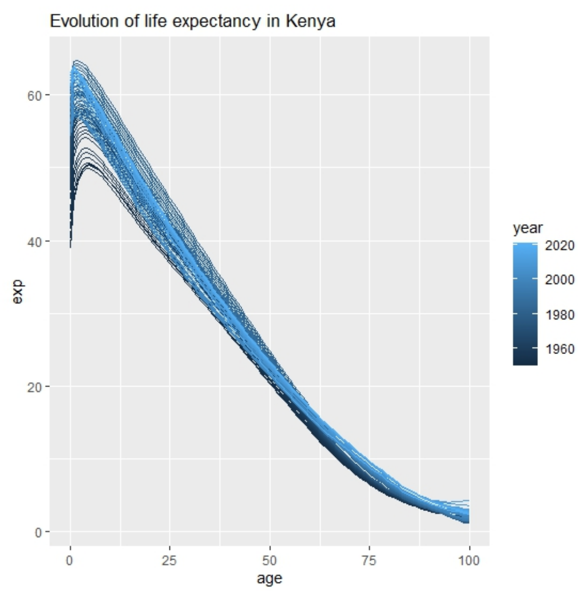
2.2.1 Life table function

Life expectancy

Life expectancy (e_x) in terms of life table function is given as;

$$e_x = \frac{T_x}{l_x} \tag{1}$$

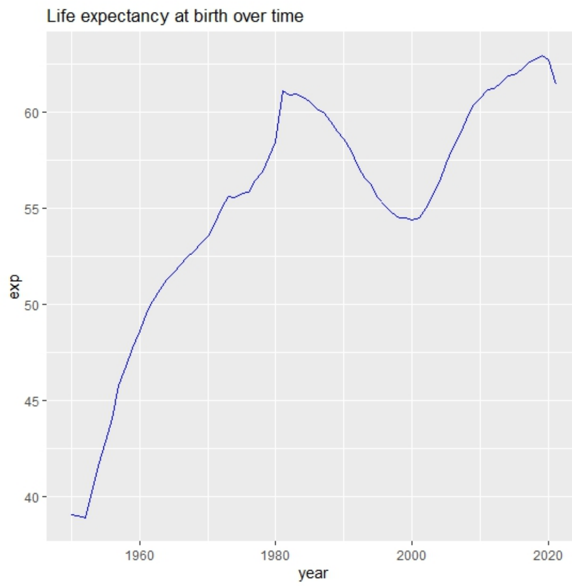
Figure 3 show the evolution of life expectancy over time in Kenya. Life expectancy at birth has increased in the recent years to more than 62 years as compared to 1950s to 1970s that was within the range of 40 to 55 years. Life expectancy at 1 year tends to be within ranges across years from 1980s.



/

Figure 3: Life expectancy evolution in Kenya plot

Life expectancy after 1 year has continually declined, with the recent years being in the middle of the bandwidth. There is evidence of a shift in life expectancy for age more 1 year and less than 62 as you approach 1980s , followed by a sharp decrease in 1990s and stabilization at the middle as the new millennium approached. Life expectancy after 62 years is higher in recent years as compared to early years. There is convergence of life expectancy in the neighborhood of



/

Figure 4: Life expectancy at birth over time

100 years across years. Figure 4 clearly show the trend in life expectancy at birth of the years. There is a continuous increase in life expectancy at birth with breaks in the trend as from 1990 to 2000 and around the Covid-19 pandemic period.

Age survival rate

Age survival curve is the plot of the share of individual expected to survive upto a certain age. This is obtained from survival function of a life table as shown figure 5 for selected years.

Its clear that less than half of the population in mid-20th century made past 50 years. In contrast, more than 80% of the population in 2020 were expected to live longer than 50 years. Individuals born in 1980 has the highest share to survive to approximately 70 years as compared to years. The survival rate of making it to adult years in the year 2000 drastically reduced as compare to 1970 and 1980.

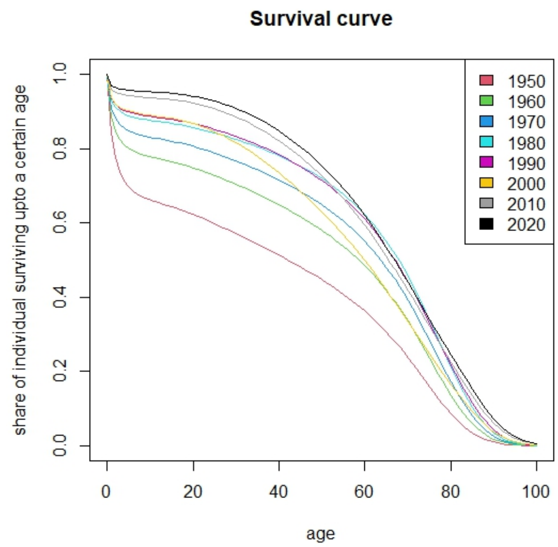


Figure 5: Survival curve

Age at death distribution

Age at death distribution is a plot of the death function of a life table as shown in figure 6. The distribution of age at death is characterized by two mode; younger mode for infant and older age mode for adult. In the early years, Kenya was higher mortality population with younger mode being greater than older age mode. This was characterized by infant mortality of 150 death per 1000 in 1950. However the trend has improved with the recent years having low younger age mode and low infant mortality of approximately 26 death per 1000 in the year 2020. The older age mode has slightly shifted over years. The gap between the younger mode and older age mode has also reduced over time, for example year 2020 younger age mode being 0.026 which is equivalent to older age mode of 0.024. This indicates a progress shift in mortality from high mortality population to intermediate status.

The dispersion of the age at death distribution has also reduced with time. There exist a heavy tail on both side of the distribution in the recent year as compared to 1950s and 1960s. The mortality risk factors tends to be wide spread across age over the recent years as compare to early years. The modal age at death corresponding to the peak of the distribution has slightly varied over time ranging from 73 to 76 years in 2020 and 1980 respectively.

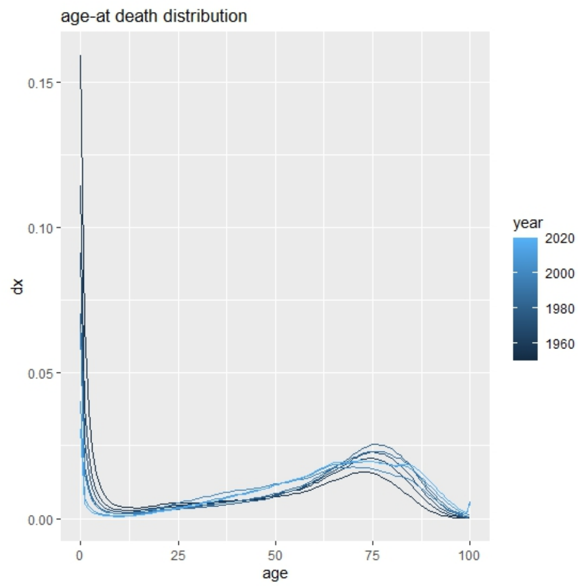


Figure 6: Age at death distribution

2.2.2 Lifespan inequality

Lifespan inequality is well explained by use of income inequality framework associated with Lorenz curve and gini index. Lorenz curve when applied to demography shows person's year lived distribution that's the cumulative person years lived share as a function of the cumulative death number share. In terms of the life table function table functions is given as (Shkolnikov et al., 2003);

$$f(x) = \frac{d(x)}{l(0)} \tag{2}$$

the cumulative death density function of person year lived x

$$F(x) = 1 - \frac{l_x}{L_0} \tag{3}$$

the cumulative share of deaths with person lived less or equal to x

$$\Phi(x) = \frac{T_0 - (T_x + xl_x)}{T_0} \tag{4}$$

is the share of total person year of life. The divergence between the diagonal and Lorenz curve indicates the variability in person year's lived. Perfect equality happens at only two end points

- $F(x) = 1, \Phi(x) = 1$ for $x = e_0$
- $F(x) = 0, \Phi(x) = 0$ for $\forall x, x \neq e_0$

Gini coefficient is a measure of absolute value mean of the inter-individual difference in age at death, divide by life expectancy (Shkolnikov et al.,2003). It normally ranges from 0 to 1. For values close to 1 indicates greater inequality. In life table notation, the gini index is given as shown in equation (5), on assumption l_0 of 1

$$G = 1 - \frac{1}{e_0} \sum_{y=0}^{\omega} l_{y+1}^2 \tag{5}$$

Gini coefficient expresses the amount of space between perfect line of equality and the lorenz curve doubled.

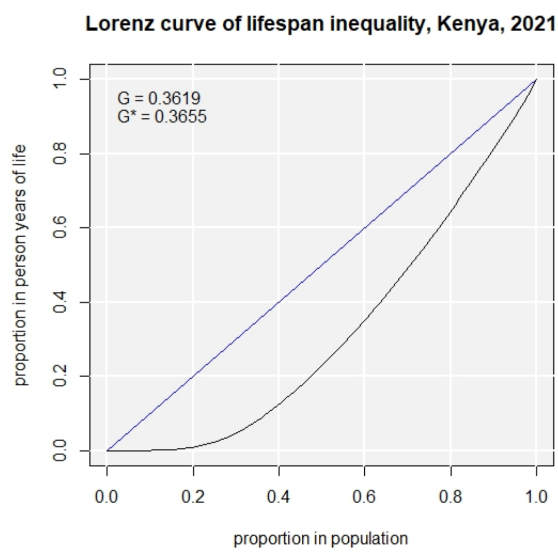


Figure 7: Lorenz curve of lifespan inequality

Considering year 2021 life table the Lorenz curve and Gini coefficient are shown in figure 7. For those longest-lived 20% person claimed 32% of the total number of life years indicated by upper part of the curve. Short-lived 40% of

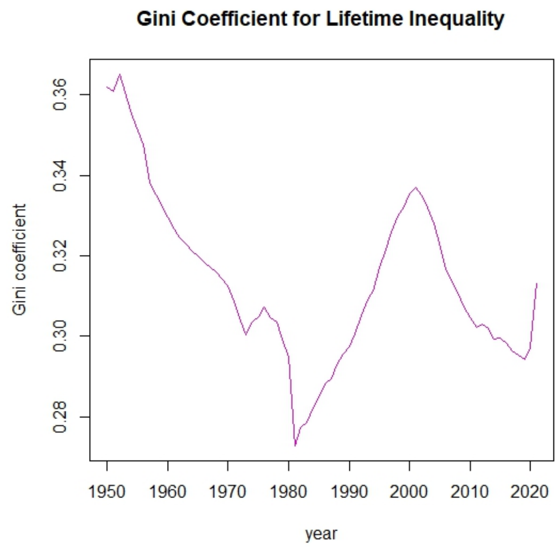


Figure 8: Gini Coefficient for Lifetime Inequality plot

Figure 9

persons claimed around 18% of all years of life shown at bottom of the curve. Gini index based on age specific mortality in 2021 is estimated to be 0.3619, indicating that person year lived from birth to death are less equally distributed. The Gini coefficient for the other year were calculated and plot as shown in figure 9. There has been a slight improvement in lifespan disparity over time with a maximum of 36 in 1950 and a minimum being 28 for 1980 when expressed as a percentage. Kenya is characterized as a population with high lifespan inequality that has not improved over time, commonly experienced in developing countries (Peltzman 2009).

2.2.3 Force of mortality

Force of mortality denoted as μ_x is the instantaneous death rate at exact age x. Its the ratio of rate of change in the number of survivors (l_x) at exact age x to the number of survivors (l_x). Mathematically given as:

$$\mu_x = \frac{-1}{l_x}\frac{dl_x}{dx} \tag{6}$$

Gompertz (1925) law of mortality is the most famous experience of the force of mortality given as;

$$\mu_x = e^{-k} e^{hx} \tag{7}$$

where h is a positive constant and k is an integration constant. Both e^{-k} and e^{h} reduces to constants B and C respectively, thus equation (7) becomes;

$$\mu_x = BC^x \tag{8}$$

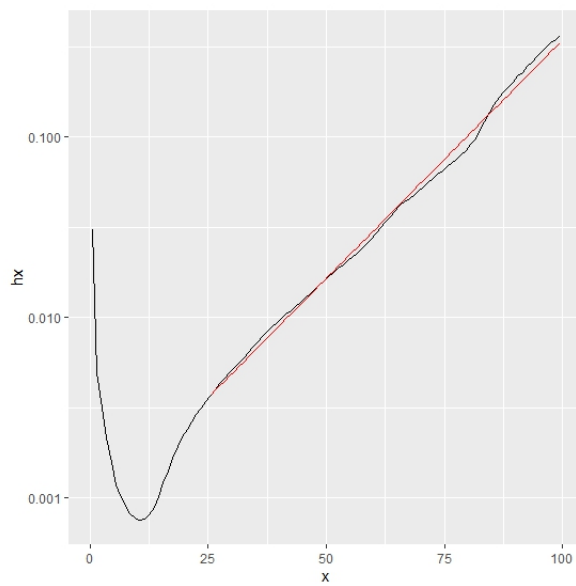


Figure 10: Force of Mortality

Table 1: Gompertz force of mortality results

a		Estimation	
	Estimate	Std. Error	P-value
(Intercept)	-5.6286523	0.0190883	2×10^{-16}
slope	0.0608969	0.0004408	2×10^{-16}
	R-Squared: 0.9962 F statistics: 1.908×10^4 on (1,73) DF p-value < 0.0001		

The estimated parameter of the Gompertz model are shown in table 1. The age from 25 years is selected to model the force of mortality as it seems to be linear as shown in figure 10. The estimated Gompertz model is statistically significant at 5% level. From the model the values of constant B and C are 0.0609 and 0.0035 respectively. The constant C states that the deaths at exact age 25 is just about 35 per ten thousand, and the slope equivalent to B indicates the death rate increases by 6.09% per year starting at age 25.

2.2.4 Malthusian parameter

Malthusian parameter also known as the intrinsic rate of natural increase. Its defined as the rate of population increase that reproduce within discrete intervals of time and has generations that are not overlapping.

$$r = \frac{lnR_0}{G} \tag{9}$$

where $R_0 = \frac{\sum l_x m_x}{l_0}$ is the net reproductive rate, $G = \frac{\sum x l_x m_x}{R_0}$ is the adult female average age of giving birth, l_x is the number of survivors at age x and m_x is the is the female offsprings per female at an exact age x (age specific fecundity)

Table 2: intrinsic rate of increase and related function results

	value(2021)	value(2000)	value(1980)
R_0	0.7261	0.5312	0.4120
G	3128.705	2786.959	2719.373
r	-0.0001	-0.0002	-0.0003

The values of $R_0 < 1$ and $r < 0$ for the years 2021,2000 and 1980, which implies that under stable population assumption Kenya is a shrinking population. Intrinsic rate of natural increase has increased from 1980, though by a small margin. Since the age of child bearing is 15 to 49 years then mean age of child bearing is less than the mean age at death as the age at death distribution shows a tendency in the age of 70s ie $E(A) < E(A)_{deaths}$. This implies that death rate decrease with increase in r. Kenya death rate has been reducing under stable population conditions as the intrinsic rate of natural increase improves.

2.2.5 Lee - Carter model estimation

Lee et al. (1992) developed a model that expresses log of age-specific death rate as a linear function of unobserved specific period index with its parameters being dependent on age as shown below.

$$ln(m_{x,t}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t} \tag{10}$$

where α_x and β_x are age dependent parameters, κ_t is a stochastic process dependent only on year of observation and $\varepsilon_{x,t}$ is the error term.

The exponential of α_x is the general shape by age while β_x describe the rate of decline in response to change in κ ie

$$\frac{dln(m_{x,t})}{dt} = \beta_x \frac{d\kappa}{dt}$$

When

- κ is linear in time: - implies each age specific mortality changes at its own constant exponential rate
- κ approaches $-\infty$: - implies each age specific mortality rates approaches zero. There is no negative death rates.

OLS method is used to estimate the parameter α_x, β_x and κ_t for $x = 1, 2, 3, \ldots, N$ and $t = 1, 2, 3, \ldots, T$, subject to boundary conditions

$$\sum_{x=1}^{N} \beta_x = 1 \quad and \quad \sum_{t=1}^{T} \kappa_t = 0 \tag{11}$$

The estimator of α_x is given as

$$\hat{\alpha}_x = \frac{1}{T} \sum_{t=1}^{T} ln(m_{x,t}) \tag{12}$$

The other parameter estimators are then obtained by singular value decomposition (SVD) of $N \times T$ matrix M, for

$$M_{x,t} = ln(m_{x,t}) - \hat{\alpha}_x \tag{13}$$

Under SVD, the solution is

$$M = UDV^T \tag{14}$$

Then, the estimator for parameter b_x and k_t are given as:

$$\hat{\beta}_x = \frac{1}{c} U_{x,1} \quad and \quad \hat{\kappa}_t = c.D_{1,1}.V_{1,t} \tag{15}$$

where
$D_{1,1}$ is the largest singular value of M
$U_{x,1}$ is the entry value at (x,1) of U
$V_{1,t}$ is the entry value at (x,t) of V and
$c = \sum_{x=1}^{N} U_{x,1}$ satisfying the boundary conditions

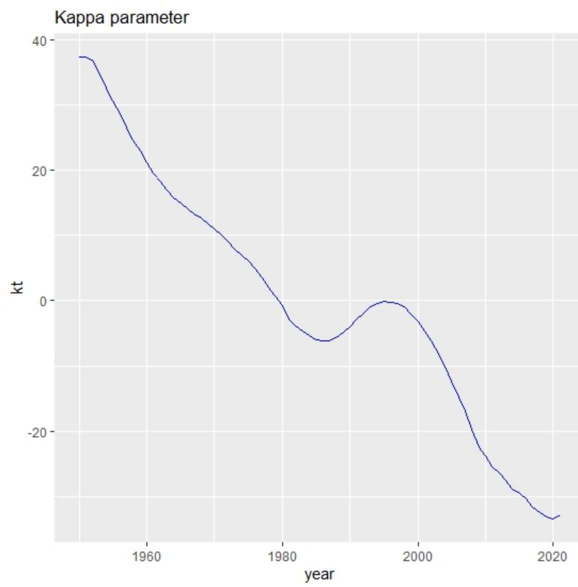
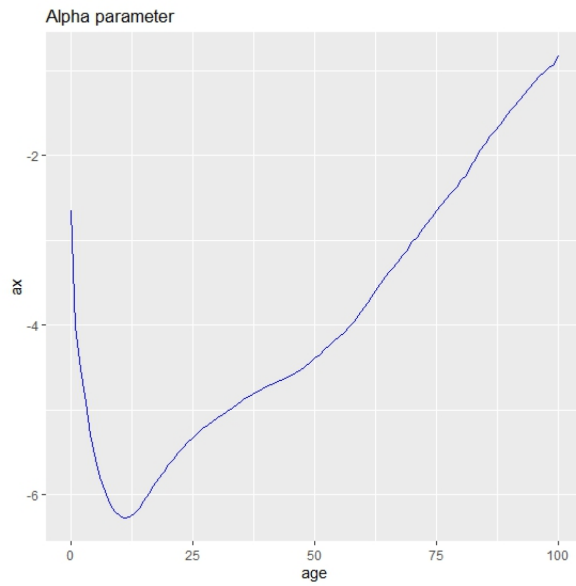
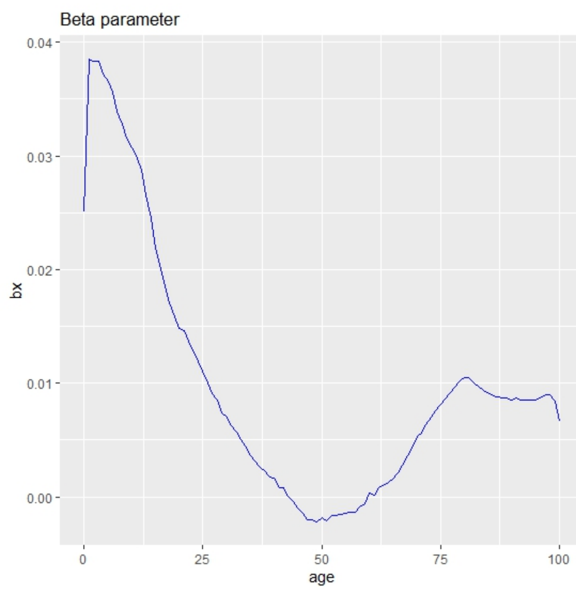


Figure 11: Kappa component plot

The Lee- Carter model estimated is statistically significant with a log-likelihood score of -84543.7 of deviance 110405.2 and 272 parameters. The model s 78.4% of the variation. General mortality pattern α is high for infant mortality, hump around ages of 25 years and linearly increases from 50 years as shown in figure 12(a). The β parameter as shown figure 12(b) is low indicating approximately uniform mortality rate change across years. For the age of 44 to 59 years, the parameter β has negative values which indicate worsening mortality in these age group over time. The mortality index decline in approximately in cubical way with two inflection points at the mid 1980s and mid 1990s respectively as shown in figure 11.



(a) Alpha plot



(b) Beta plot plot

Figure 12

2.2.6 Forecast

The estimated Lee-Carter model is used to forecast life expectancy. The Lee-Carter model set κ parameter into a time series, which then use ARIMA process to project forward time series value.

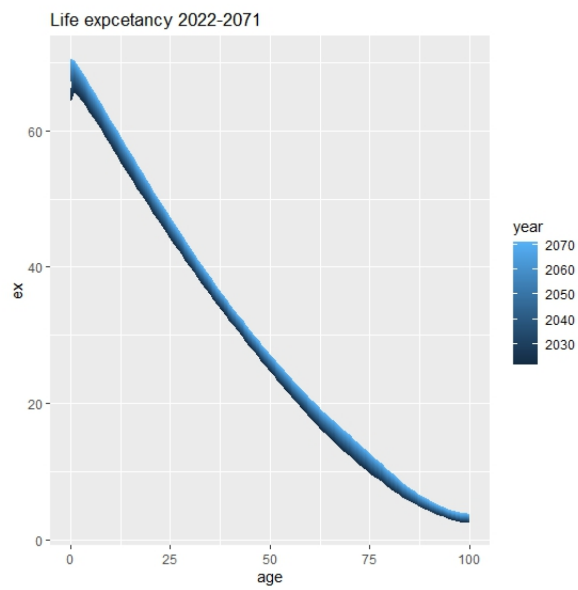


Figure 13: Expected number of years left plot

Table 3: Gompertz force of mortality results

ERROR MEASURES BASED ON MORTALITY RATES			
Averages across ages			
ME	MSE	MPE	MAPE
-0.00019	0.00060	0.02092	0.11080
Averages across years			
IE	ISE	IPE	IAPE
-0.01537	0.04937	2.08505	11.06310

ME- mean error; MSE - mean square error; MPE- mean percent error; MAPE- mean absolute percent error; IE - integrated error; ISE- integrated square error; IPE- Integrated percent error and IAPE - integrated absolute percent error

The fitted process is ARIMA(1,1,0). To assess the accuracy of forecast, average

error measure of mortality rate both across age and year are considered as shown in table 3. Mean average percent error (MAPE) is used to tell how far off the predicted values are on average. On average across age, the predictions are less than 1% off the actual values. The average across years is slightly higher, at 11% off the actual values.
Life expectancy from year 2022 to 2071 is shown in figure 13. Generally life expectancy at any given age is projected to increase with time. Life expectancy at birth in 2022, 2023 and 2024 are 64.4, 64.5 and 64.7 respectively. Its projected that as Kenya attains a middle class developed country by 2030, life expectancy at birth would be 65.6 years.

3 Discussion and conclusion

The study sort to analyses the trend and model mortality by considering life table functions, lifespan inequality, force of mortality and expected life expectancy at any given age. The life table functions demonstrated clear pattern in the characteristic of the life expectancy evolution, age distribution of death and modal age trend, and proportion of the population able to survive to a given age. The Lorenz curve and Gini coefficient satisfactory exhibited the underlying pattern in lifespan inequality. The Gompertz and Lee-Carter model estimated were statistically significant at 5% confidence level.

This study provides evidence to conclude that: life expectancy at an exact given age has increased; infant mortality has continuously reduced to approximately 26 per 1000 in 2020; Covid-19 had negative effect on mortality characteristics resulting to reduced life expectancy,survival rate, age at death distribution parameters and increase in lifespan inequality; year 2000 is characterized by extreme low mortality characteristics ie sharp decline in survival rates, sharp decrease in life expectancy, wide dispersion of age at death distribution and high lifespan inequality; year 1980 has best combination of mortality characteristic over the entire period; Person year lived from birth to death are less equally distributed and a slight improvement with time; The country has shifted from high mortality population, at an intermediate stage; modal age at death has slightly reduced to 73 years in 2020; increase in mortality risk factors across age; the force of mortality in year 2021 as from exact age 25 is linear with an annual rate of increase of 6.09% from 35 per ten thousands; Intrinsic rate of natural increase has been improving though less than zero, and mortality rate is decreasing under stable population assumptions; there is a worsening mortality rate at the age of

44 to 59 years; there is approximately uniform low mortality rate of change at a given year over time; and Expected life expectancy at a given age is projected to increase, with life expectancy at birth in 2030 and 2071 being 65.6 and 70.5 years respectively.

We recommend that policy makers to develop measure and framework to address the worsening mortality at the age of 44 to 59, the spread of health risk factors across age and lifespan inequality. In addition, the child mortality measures being implemented be strengthen further as the population shift in mortality status. For the period 1990 to 2020, mortality characteristics were worsening, We recommend future research to be undertaken to understand the drive force that led to such scenarios.

References

[1] Booth, H., Maindonald, J., & Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. Population studies, 56(3), 325-336.

[2] Booth, H., & Tickle, L. (2003). The future aged: new projections of Australia.'s elderly population. *Australasian Journal on Ageing*, 22(4), 196-202.

[3] Canudas-Romo, V. (2008). The modal age at death and the shifting mortality hypothesis. *Demographic Research*, 19, 1179-1204.

[4] Canudas-Romo, V. (2010). Three measures of longevity: Time trends and record values. *Demography*, 47, 299-312.

[5] Carter, L. R., & Prskawetz, A. (2001). Examining structural shifts in mortality using the Lee-Carter method. Methoden und Ziele, 39.

[6] Chavhan, R., & Shinde, R. (2016). Modeling and forecasting mortality using the Lee-Carter model for Indian population based on decade-wise data. *Sri Lankan Journal of Applied Statistics*, 17(1).

[7] Cheung, S. L. K., Robine, J. M., Tu, E. J. C., & Caselli, G. (2005). Three dimensions of the survival curve: Horizontalization, verticalization, and longevity extension. *Demography*, 42(2), 243-258.

[8] Debón, A., Montes, F., & Puig, F. (2008). Modelling and forecasting mortality in Spain. *European Journal of Operational Research*, 189(3), 624-637.

[9] Ebeling, M., Rau, R., & Baudisch, A. (2018). Rectangularization of the survival curve reconsidered: The maximum inner rectangle approach. Population Studies, 72(3), 369-379.

[10] Fazle Rabbi, A. M., & Khan, H. T. (2022). Stochastic mortality forecasts for Bangladesh. Plos One, 17(11), e0276966.

[11] Giorgi, G. M., & Gigliarano, C. (2017). The Gini concentration index: a review of the inference literature. *Journal of Economic Surveys*, 31(4), 1130-1148.

[12] Horiuchi, S., Ouellette, N., Cheung, S. L. K., & Robine, J. M. (2013). Modal age at death: lifespan indicator in the era of longevity extension. *Vienna Yearbook of Population Research*, 37-69.

[13] Ibrahim, N. S. M., Lazam, N. M., & Shair, S. N. (2021, July). Forecasting Malaysian mortality rates using the Lee-Carter model with fitting period variants. *In Journal of Physics: Conference Series (Vol. 1988, No. 1, p. 012103).* IOP Publishing.

[14] Keilman, N. (1998). How accurate are the United Nations world population projections?. *Population and Development Review*, 24, 15-41.

[15] Lee, R. D. and Carter, L. R. (1992): Modeling and Forecasting the Time Series of U.S. Mortality. *Journal of the American Statistical Association*, 87(419): 659–671.

[16] Lee, R.D., & Rofman, R. (1994). Modeling and forecasting mortality in Chile. Natas 22 (59), 182–313.

[17] Lundström, H., & Qvist, J. (2004). Mortality forecasting and trend shifts: An application of the Lee–Carter model to Swedish mortality data. International Statistical Review, 72(1), 37-50.

[18] Roser, M., Ortiz-Ospina, E., & Ritchie, H. (2013). Life expectancy. Our world in data.

[19] Shang, H. L., & Haberman, S. (2020). Forecasting age distribution of death counts: An application to annuity pricing. *Annals of Actuarial Science*, 14(1), 150-169.

[20] Shkolnikov, V. M., Andreev, E. E., & Begun, A. Z. (2003). Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples. Demographic Research, 8, 305-358.

[21] Peltzman, S. (2009). Mortality inequality. Journal of Economic Perspectives, 23(4), 175-190.

[22] Tuljapurkar, S., Li, N., & Boe, C. (2000). A universal pattern of mortality decline in the G7 countries. Nature, 405(6788), 789-792.

[23] United Nation (2022). World population prospect *https://population.un.org/wpp/Download/Standard/Mortality/*

[24] Vaupel, J. W., Zhang, Z., & van Raalte, A. A. (2011). Life expectancy and disparity: an international comparison of life table data. *BMJ open*, 1(1), e000128.

[25] Zafeiris, K. N. (2023). Comparing the Mortality Regimes in 39 Populations. *In Quantitative Demography and Health Estimates: Healthy Life Expectancy, Templates for Direct Estimates from Life Tables and other Applications* (pp. 187-204). Cham: Springer Nature Switzerland