UiO : Department of Mathematics
University of Oslo

Modeling and Forecasting Norwegian mortality for pension and life insurance purposes

Thomas Løland

Master's Thesis, Spring 2022

This master's thesis is submitted under the master's programme *Stochastic Modelling, Statistics and Risk Analysis*, with programme option *Finance, Insurance and Risk*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 30 credits.

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Abstract

Forecasting mortality is crucial in life insurance sciences. The accuracy of such forecasts is essential to compute the present value of future promised cash flows in life and pension insurance. The Lee-Carter model is a popular demographic model used for estimation and prediction of mortality rates. In this thesis the standard Lee-Carter model will be used on Norwegian Mortality rates from the Human Mortality Database. After presenting the model in detail, its goodness of fit and prediction capabilities will be assessed. The forecasts will be compared to the proposed mortality rates by the financial supervisory authority of Norway in their document K2013. Based on the forecasts of the Lee-Carter model, the present values and mathematical reserves of both pension and endowment insurances will be calculated and analyzed.

Acknowledgements

This master's thesis is the final part of my masters degree in *Finance, Insurance and Risk* at the University of Oslo, and therefore marks the end of years of studying. It is already with a certain nostalgia I look back at ups and downs over the course of this period of my life.

I would like to take the opportunity to thank the Department of Mathematics and all its lecturers and faculty members for all the great experiences over course of my studies. Among them, I especially thank my supervisor David Ruiz Baños for his inspiring lectures and all the good advise through the writing process.

I owe an tremendous debt to my friends and family for supporting my academic efforts. But most of all I owe my girlfriend for always shining a light on my shadow of doubt. Her support has made all the difference, and for that I am very thankful.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Forecasting mortality rates is crucial for insurers and policy makers to assess the cost of increasing life expectancy. The population in a well-developed country like Norway live longer and longer due to medical advancements, fewer work incidents, less risky lives and other factors. The importance of capturing this trend in life expectancy is essential to both private and public pension insurers. To calculate the present value of policies, the models of future mortality need to able to predict with high accuracy what the future mortality rates are going to be.

Since the first attempts to measure mortality, a lot of progress has been made: since the ground-breaking Gompertz's *law of mortality* in 1825[Gom25] many different models have been proposed. One of these being the Lee-Carter method.[LC92] This two-dimensional model quickly became widely popular and used for various applications. Various extensions to the model have been proposed over the years, but the basic model first proposed by Lee and Carter in 1992 to forecast mortality in the United States has shown to be a simple yet robust method. This thesis sets out to model Norwegian mortality based on the original configuration of Lee and Carter. Based upon historical Norwegian death counts and exposures to risk from the Human Mortality Database, log mortality rates will be calculated. Following the forecasting procedure of the Lee-Carter model, future mortality rates will then be predicted.

A goal of this thesis is to clearly lay out the different steps necessary to the procedure. The resulting estimations will be assessed according to how good they fit the actual data. To evaluate the prediction power of the Lee-Carter model, Norwegian mortality data will be divided into a test and training set.

Future mortality forecasts for years $2021 - 2060$ will be compared to future mortality rates proposed by the financial supervisory authority of Norway in their well-known document K2013, to see how well the two models match each other. Thereby, the forecasted Lee-Carter mortality rates will be used in both an endowment and a pension insurance scenario to compute the present values and mathematical reserves of such policies. Conclusions will be made as to what these findings suggest.

1.1 Outline

The rest of the thesis is organized as follows:

Chapter 2 This chapter introduces the basic concepts of mortality modeling as well as the mathematical tools necessary to the Lee-Carter model

Chapter 3 This chapter starts with an explanation of Gompertz's law of mortality before the Lee-Carter model will be detailed in full.

Chapter 4 In this chapter the Lee-Carter model is applied to the Norwegian mortality rates. This includes estimation, forecasting and analysis.

Chapter 5 This chapter starts with the theory behind calculating present values of policies and calculation of mathematical reserves. Then this will be applied in practice based on forecasted Norwegian mortality rates from chapter 4.

Chapter 6 The conlusion of the thesis as well as future work will be laid out in this chapter.

Appendix A Proofs

Appendix B R-code used for calculating and displaying the results of the thesis.

CHAPTER 2

Prerequisites

In this chapter, the prerequisites for our intended mortality modeling will be detailed. In the first section, all the necessary concepts and definitions of actuarial sciences will be explained, followed by a section with some basic concepts required for modeling the Lee-Carter model.

2.1 Basic Actuarial Concepts

The modeling of mortality rates requires some basic actuarial concepts. These will be explained in the subsequent subsections.

Measuring Mortality

"Mortality" refers to the number of deaths for a specific area in a specific period of time, either cause-specific or all-cause. The time-variable, which we denote T_x, is a non-negative random variable, distributed on $[0, \infty)$. For a given individual (x) alive at exact age x, T_x represents the remaining *future lifetime*. Consequently, we have that $x + T_x$ is the random variable of age-at-death for individual (x).

Now, as described in the book by Dickson, Hardy and Waters([DHW19]), we let F_x be the distribution function of T_x, defined as such:

$$F_x(t) = P[T_x \leq t],\ t \geq 0. \tag{2.1}$$

This is referred to as the *lifetime distribution* from age x. For a given t and x, this is the *mortality rate*, i.e. it is probability that an individual with exact age x dies between time t and $t+1$. However, in many cases, we are interested in the probability of death occurring later than some specified time t. This probability is called the *survival function*, and is denoted S_x

$$S_x = 1 - F_x(t) = P[T_x > t]. \tag{2.2}$$

From basic probability theory, it is trivial that this function is the complement of the distribution function F_x.

Alternative notation

It is common to stumble upon different actuarial notations. In actuarial sciences, the lifetime distribution in 2.1 is often denoted as[DHW19]

$$ {}_tq_x = F_x(t) = 1 - S_x(t), \tag{2.3}$$

in which the t notation is usually dropped when we let $t = 1$, in which case q_x becomes the yearly mortality rate for an individual aged exactly x. Furthermore, the survival function is usually denoted[DHW19]

$$ {}_tp_x = S_x(t), \tag{2.4}$$

where, as in equation 2.3, the t is skipped from the notation for $t = 1$.

Hazard rate (Force of Mortality)

The Hazard Rate, also called Force of Mortality, transition intensity or failure rate (depending on the field of study) is the probability that an individual or component dies or fails between the times $x + t$ and $x + t + \Delta t$ where Δt decreases to zero ([MRC18]). I.e. it is the *instantaneous* rate of death/failure for an individual or component at time $x + s$. In actuarial sciences, given an individual aged $x+t$, we let $\mu_{x,t}$ denote the Hazard Rate, and with our notation, as defined in the book by Macdonald et al., it is given as[MRC18]

Definition 2.1.1. The hazard rate or force of mortality at age $x + t$ associated with the random lifetime T_x is:

$$\mu_{x,t} = \lim_{\Delta t \to 0^+} \frac{P[T_x \leq t + \Delta t \mid T_x > t]}{\Delta t}, \tag{2.5}$$

Since μ_x is non-negative, it has the following properties:

1. $\mu_x \geq 0$ for all $x \geq 0$
2. $\int_0^\infty \mu_x dx0 = \infty$

The hazard rate has an important relationships with the survival function $S_x(t)$. It can be shown that S_x can be written as

$$S_x(t) = \exp\left(-\int_0^t \mu_{x+s} ds \right). \tag{2.6}$$

The proof of this relationship is given in appendix A.

The Central Rate of Mortality

The Central Rate of Mortality is defined as the average incidence of deaths in a population aged x in a particular time period; i.e. it is found by dividing the average number of deaths aged exactly x during the time period in question with the average number of individuals alive at that age group during the period. Denoted m_x, the Central Rate of Mortality as defined in the book of Macdonald et al. is given as follows[MRC18]

Definition 2.1.2. The Central Rate of Mortality at age x, denoted m_x is defined as

$$m_x = \frac{q_x}{\int_0^1 {}_tp_x dt}, \tag{2.7}$$

where q_x is the probability that an individual aged exactly x will die before reaching age $x+1$, and ${}_tp_x$ is the survival function as defined in equation (2.4).

Furthermore, by using the relationship detailed in equation 3.18 in the book by Macdonald et al. which states that

$${}_tq_x = \int_0^t {}_sp_x \mu_{x+s} ds,$$

we can write equation (2.7) as follows:

$$m_x = \frac{\int_0^1 {}_tp_x \mu_{x+t} dt}{\int_0^1 {}_tp_x}. \tag{2.8}$$

The Crude Hazard Rate

The Crude Hazard Rate is defined as the number of deaths occurring among a population during a given time period divided by the number of years lived by that population over the same time period. For the age interval x to $x+1$, we denote the Crude Hazard Rate by $\hat{r}_x$, which is defined to be[MRC18]:

$$\hat{r}_x = \hat{\mu}_{x+s} = \frac{d_x}{E_x}.$$

In the above equation, $\hat{\mu}_{x+s}$ is the estimated hazard rate from equation (2.5) based on historical data d_x and E_x. d_x denotes the number of deaths among the population aged exactly x over a certain time period, while E_x denotes "exposed to risk" in the same population aged x over the same time period, which is exposure based on "time lived." As long as the hazard rate is not rapidly changing, the s can reasonably be set to 1/2, thus making the Crude Hazard Rate an estimate for $\mu_{x+1/2}$([MRC18])

$$\hat{r}_x = \hat{\mu}_{x+1/2} = \frac{d_x}{E_x}. \tag{2.9}$$

2.2 Mathematical tools

In the subsequent sections the mathematical tools and results required to perform the modeling of mortality in the Lee-Carter model will be introduced.

The Singular Value Decomposition (SVD)

The Singular Value Decomposition (SVD) is a way to factorize real or complex matrices. It is useful in a wide range of science applications, such as image processing, gene expression analysis and topographical analysis (see e.g. [Sad12] [Sma94], [WRR02]). For these purposes, the power of SVD lies mainly in compressing data. In our application, the Singular Value Decomposition is used to estimate the parameters in the Lee-Carter model (section 3.3). Given a real or complex $m \times n$-matrix, the SVD is defined as[TB97](p. 28-29))

Definition 2.2.1. The Singular Value Decomposition of a $m \times n$ matrix A is

$$A = U\Sigma V^T, \tag{2.10}$$

where U is a unitary $m \times m$ matrix, Σ is a diagonal $m \times n$ matrix with strictly positive real numbers on the diagonal, and V^T is a $n \times n$ unitary matrix, where V^T denotes the matrix transposed of V.

By matrix multiplication of equation 2.10 it is trivial to see that the SVD can be written as

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T,$$

where r is the rank of A, the u_i's and v_i^T's denote the diagonal elements of U and V^t respectively (up until $r = rank(M)$), and the σ_i's are the diagonal elements of Σ. The columns of U and V are called the left and right *singular vectors*, while the diagonal entries of Σ are called the *singular values* of A, with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r \geq 0$, i.e. non-negative and decreasing in i.

The Autoregressive Integrated Moving Average model (ARIMA)

The Autoregressive Integrated Moving Average model (ARIMA for short) is used to in time series analysis for forecasting purposes. It is one of the most widely used methods in this regard([HA21]). Later in this thesis, ARIMA will used to forecast the time series k_t in the Lee-Carter model. The acronym ARIMA can be broken down into three constituent parts of the procedure: The autoregregression, AR, part tells us that the regression of the time-dependent variable is performed on its previous value; the "integrated", I, of the acronym specifies that the data points are the difference between the current and preceding values of the time series on which the analysis is performed; while the MA part refers to the regression error being represented as a (weighted) moving average of the past forecasting errors. For a time series X_t, the ARIMA(p, d, q)-model can be written as

$$X_t^{'} = \delta + \alpha_1 X_{t-1}^{'} + \cdots + \alpha_p X_{t-p}^{'} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t,$$

where p denotes the order of the autoregressive part, d refers to the degree of first differencing and q is the order of the moving average part([HA21]), while δ is a contant, the α_i $(i = 1 \ldots p)$ are the parameters of the autoregressive part of the model, the θ_i $(i = 1 \ldots q)$ are the parameters of the moving average part and the ϵ_t are the error terms. $X_t^{'}$ is the differenced data, which for $d = 1$ is $X_t^{'} = X_t - X_{t-1}$, while for a second-order differencing $(d = 2)$ $X_t^t = X_t^* - X_{t-1}^*$, where X_t^* is the first-order differencing of X_t.

Since the parameters p, d and q refers to the different parts of the ARIMA-model, setting one or more of these to zero reduces the complexity of the procedure. E.g. by setting d and q to zero (ARIMA$(p, 0, 0)$), we eliminate the differencing and moving average part and are left with a autoregression model, conveniently acronomized as a AR-model. Likewise, for ARIMA$(0, 0, q)$, we are only left with the moving average part of the procedure, simply called a MA-model. The special case of the ARIMA$(0, 1, 0)$, also called the $I(1)$-model, is of interest to our goals in this thesis. Observe that this model, with constant $\delta = 0$ is simply the random walk([HA21]):

Definition 2.2.2. The random walk is defined as

$$\begin{gathered} ARIMA(0,1,0) = I(1) = X_t = X_{t-1} + \epsilon_t, \\ \epsilon_t \sim \mathcal{N}(0, \sigma^2). \end{gathered}$$

Where X_t is a real-numbered time series with integer index t. The ϵ_t is the i.i.d error term.

Thus, the random walk is a special case of the ARIMA-model. If, however, we allow the constant δ to be non-zero, we get a random walk with a constant drift δ:

Definition 2.2.3. The random walk with drift is defined as

$$\begin{gathered} ARIMA_\delta(0,1,0) = I_d(1) = X_t = X_{t-1} + \delta + \epsilon_t, \qquad (2.11) \\ \epsilon_t \sim \mathcal{N}(0, \sigma^2). \end{gathered}$$

Where X_t is a real-numbered time series with integer index t, and δ is a constant drift parameter. The ϵ_t is the i.i.d error term.

The constant δ is added if the trend of the random walk is expected to either increase or decrease over time. If it is negative the trend will be decreasing, if it is positive the trend is increasing. For a random walk to be truly random, it is a requisite that the error term is $\mathcal{N}(0, \sigma^2)$ distributed.

CHAPTER 3

Mortality models

In this section we go through the Gompertz's Law of mortality and the Lee-Carter model in detail.

3.1 Gompertz's Law of Mortality

In 1825 the British actuary Benjamin Gompertz wrote a letter to the Philosophical Transactions of the Royal Society of London[Gom25] in which he detailed his findings on the behavior of human mortality and expressed a new method of determining the value of life contingencies. This land-mark paper pioneered the mathematical study of mortality. Among his observations, Gompertz stated his law of mortality, a *parametric model* that explains the intensity of human mortality as a parametric function of age x.

Definition 3.1.1. Gompertz's law of mortality states

$$\mu_x = \mu(x; a, c) = ac^x, \tag{3.1}$$

where a is a constant quantity denoting the mortality at the initial age $x = 0$, and c is a constant denoting the exponentially increasing rate of mortality over age x. Note that the initial age is not necessarily the biological age 0, but rather the initial age of analysis; in this case, the x should be replaced by $x - x_0$, where x denotes the age and x_0 denotes the starting age of analysis, ensuring zero value at $x = x_0$. For our application, however, it represents biological age 0.

Now, taking the logarithm of equation (3.1), we get

$$\log(\mu_x) = \log(a) + x\log(c), \tag{3.2}$$

which is a linear regression model with $\beta_0 = \log(a)$ and $\beta_1 = \log(c)$, linear in age parameter x. We demonstrate Gompertz model by fitting this linear regression in R on Norwegian crude mortality rates as explained in section 2.1.6, equation 2.9 ($\mu_{x+1/2} = \frac{d_x}{E_x}$), with the data detailed in chapter 4.

Figure 3.1 shows the fitted $\log(d_x/E_x)$ and the observed $\log(d_x/E_x)$ for ages 40 to 90 in the year 2000. For this age-span the Gompertz law of mortality gives a good approximation to the empirical data. The data follows a linear trend, highlighting the power of Gompertz's observation of human mortality.

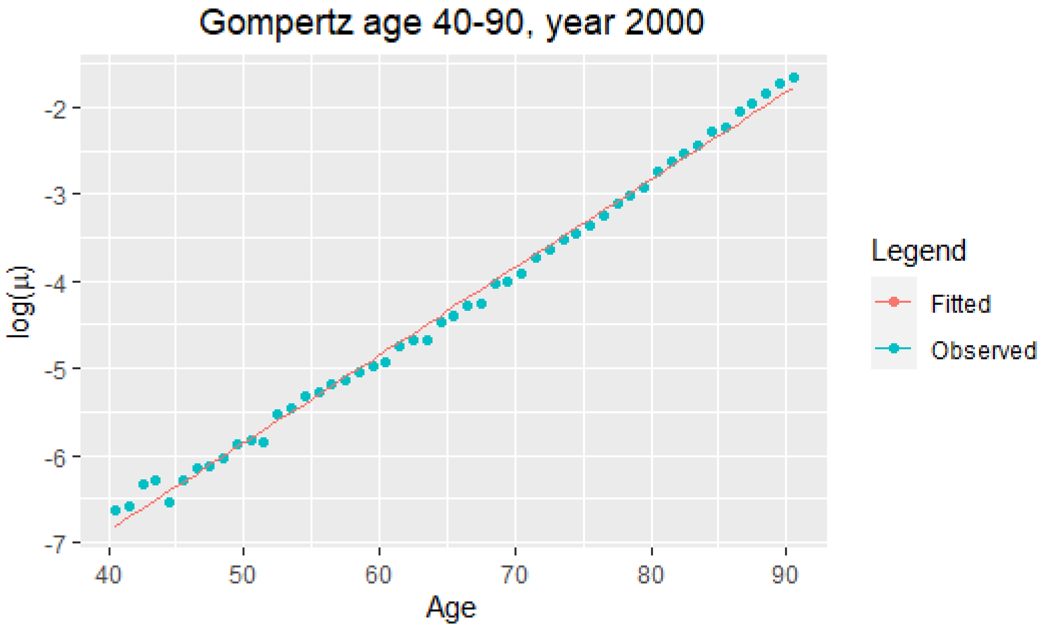

Figure 3.1: Gompertz fitted $\log(d_x/E_x)$ vs observed $\log(d_x/E_x)$. Data: Norwegian population, ages 40 to 90 in year 2000.

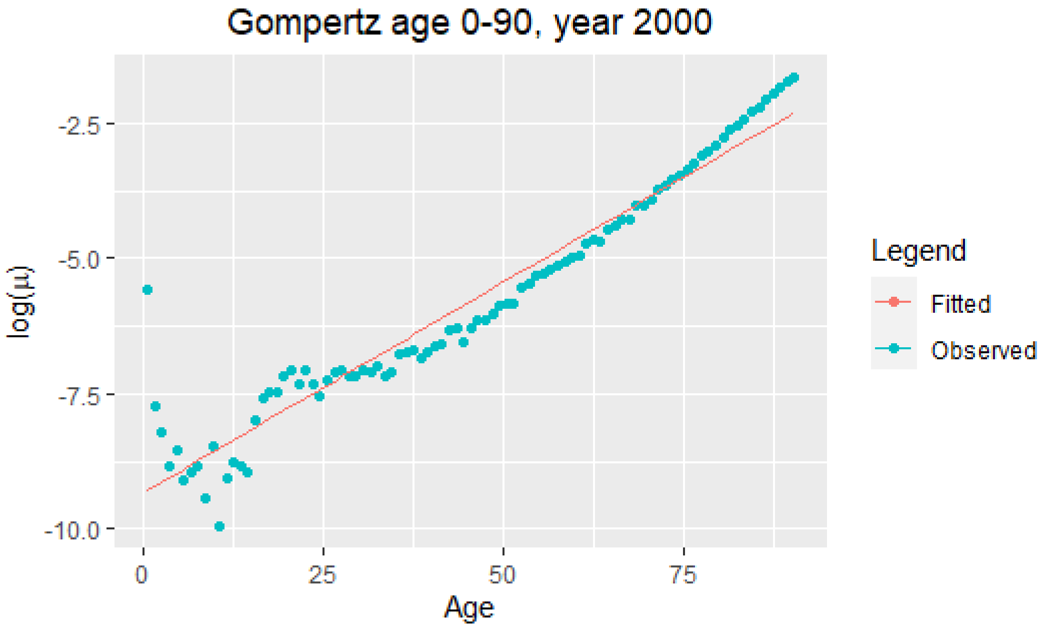

Figure 3.2: Gompertz fitted $\log(d_x/E_x)$ vs observed $log(d_x/E_x)$. Data: Norwegian population, ages 0 to 90 in year 2000.

Figure 3.2 displays fitted $\log(d_x/E_x)$ together with the observed $\log(d_x/E_x)$ for ages 0 to 90 in the year 2000. For this extended age-span the Gompertz model gives a worse fit: We see that it fails to capture the non-linearity of the data at the younger ages, especially the infant mortality. There is evidently a flaw in the assumption of linearity for this age-span.

3.2 The Lee-Carter model (LC)

The Lee-Carter model was first introduced in 1992 by Ronald D. Lee and Lawrence R. Carter in the Journal of the American Statistical Association[LC92]. Originally developed for all-cause mortality data in the United States in the period 1933-1987, it is now renowned as the leading method for modeling human mortality. The model describes a log-transform of empirical mortality rates as the sum of two age-specific parameters and one time-varying parameter.

Definition 3.2.1. The Lee-Carter model is defined as

$$\begin{aligned} \log(m_{x,t}) &= a_x + b_x k_t + \epsilon_{x,t}, \qquad (3.3) \\ \epsilon_{x,t} &\sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

where $m_{x,t}$ denotes the central rate of mortality in equation 2.7 at age x in year t, a_x is an age-specific parameter describing the general pattern of mortality at age x, b_x is an age-specific parameter describing the relative speed of change in mortality at age x, k_t denotes a time-varying mortality index and $\epsilon_{x,t}$ is the associated error term, which is assumed to be normal distributed.

Note that the parametrization in equation (3.3) is invariant under these linear transformations:

$$(a_x, b_x, k_t) \rightarrow (a_x + cb_x, \frac{b_x}{d}, d(k_t - c)) \qquad (3.4)$$

for any constants c and d, $d \neq 0$.

Therefore the solution is not unique. To obtain a unique solution, Lee and Carter imposed the following parameter constraints, which we also will be using in this study:

$$\sum_{x=1}^{X} b_x = 1, \qquad (3.5)$$

$$\sum_{t=1}^{T} k_t = 0, \qquad (3.6)$$

where X is the total number of age groups and T is the number of time periods. In our case, since we will group age and time in one year increments, these integers respectively represent the last age and last year of analysis.

3.3 Parameter Estimation in LC

From constraint $\sum_t^T k_t = 0$ (equation 3.6) we get

$$\sum_t^T k_t = 0 \implies \frac{1}{T}\sum_t^T \ln(m_{x,t}) = \hat{a}_x, \tag{3.7}$$

thus a_x is simply the empirical average over time of $\log(m_{x,t})$ in age group x. The parameters b_x and k_t can be estimated via maximum likelihood. This procedure can be quite tedious and as Lee and Carter pointed out, there is an easier solution to find the optima: the parameters can easily be found via the singular value decomposition (SVD) of the matrix of centered age profiles. We call this matrix M:

$$M = \log(m_{x,t}) - \hat{a}_x, \quad t = \tag{3.8}$$

By taking the SVD of this matrix M we obtain the least square solutions. From the definition of SVD we get2.10

$$\text{SVD}(M) = U\Sigma V^T = \sigma_1 U_{x,1}V_{t,1}^T + \sigma_2 U_{x,2}V_{t,2}^T + ... + \sigma_1 U_{x,k}V_{t,k}^T, \tag{3.9}$$

where V^T denotes the matrix transpose of V and $k = rank(M)$. Furthermore we have the singular values

$$\sigma_i, \quad i = 1, 2, ..., k,$$

and the corresponding singular vectors

$$U_{x,k}, V_{t,k}^T, \quad i = 1, 2, ..., k.$$

Our estimates of $\hat{b}_x$ and $\hat{k}_t$ are then given by[KAM16]

$$\begin{aligned}\hat{b}_x &= U_{x,1},\\ \hat{k}_t &= \sigma_1 V_{t,1}^T.\end{aligned}$$

hence our fitted Lee-Carter model becomes

$$\log(\hat{m}_{x,t}) = \hat{a}_x + \hat{b}_x\hat{k}_t = \hat{a}_x + \sigma_1 U_{x,1}V_{t,1}^T, \tag{3.10}$$

where $\ln(\hat{m}_{x,t})$ is the estimation of empirical log-mortality rates for each age x and year t and $\hat{a}_x$ is, as previously stated, the empirical average over time of the log mortality rate for age x (equation 3.7). We denote by $\hat{M}$ its matrix form, which written out looks like this

$$\hat{M}_{x,t} = \begin{bmatrix} \hat{a}_1 + \sigma_1 U_{1,1}V_{1,1}^T & \cdots & \sigma_1 U_{1,1}V_{T,1}^T \\ \vdots & \ddots & \vdots \\ \hat{a}_X + \sigma_1 U_{X,1}V_{1,1}^T & \cdots & \hat{a}_1 + \sigma_1 U_{X,1}V_{T,1}^T \end{bmatrix} \tag{3.11}$$

3.4 Forecasting in LC

After obtaining the estimates, the second step of the Lee Carter model is to produce forecasts up until a future time point. In order to do this, we need to make future predictions of our time-varying variable k_t. This is done using an *Autoregressive Integrated Moving Average* (ARIMA) model. The specification used in Lee and Carter [LC92] is that of a random walk with drift. Depending on the data set, other specifications might be preferable. However, in this study we will implement the random walk with drift. This model is as follows:

Using the equation for a random walk with drift as specified in equation 2.11, we get our forecasted estimator $\hat{k_t}$

$$\begin{aligned} \hat{k_t} &= \widehat{k_{t-1}} + \hat{\delta} + \epsilon_t, \\ \epsilon_t &\sim \mathcal{N}(0, \sigma^2), \end{aligned} \tag{3.12}$$

where $\hat{\delta}$ is the maximum likelihood estimate of the drift, which depends only on the first and last data points of the estimate $\hat{k}_t$. It is calculated as[GK08]

$$\hat{\delta} = \frac{\hat{k_T} - \hat{k_1}}{T - 1}. \tag{3.13}$$

This drift expression is used to calculate all the future values of $\hat{k}$. First we forecast two periods ahead and expand equation 3.12 by substituting $\hat{k}_{t-1}$ with its definition from the same equation and get

$$\begin{aligned} \hat{k}_t &= \hat{k}_{t-1} + \hat{\delta} + \epsilon_t \\ &= (\hat{k}_{t-2} + \hat{\delta} + \epsilon_{t-1}) + \hat{\delta} + \epsilon t \\ &= \hat{k}_{t-2} + 2\hat{\delta} + (\epsilon_{t-1} + \epsilon_t) \end{aligned}$$

We now iterate over this procedure to obtain the forecast for $\hat{k}_t$ expressed as

$$\hat{k}_t = \hat{k}_t + \hat{\delta} t + \sum_{i=1}^{t} \epsilon_{T+i-1}. \tag{3.14}$$

Since we forecast forward in time, we have that $t > T$ in our expression. By taking the expectation and variance of equation 3.14 we obtain

$$\begin{aligned} E[\hat{k}_t|\hat{k}_1, ..., \hat{k}_{t-1}] &= \hat{k}_T + \hat{\delta} t, \\ Var[\hat{k}_t|\hat{k}_1, ..., \hat{k}_{t-1}] &= t\sigma^2 \end{aligned}$$

Thus, the forecast point estimate is a function of t that follows a straight line. Therefore, forecasting $\hat{k}_t$ is merely extrapolating a straight line through the first and last data points. Furthermore, we see that both the expectation and variance of our time-evolving variable $\hat{k}_t$ depends on time. By the assumption of decreasing mortality rates over time, the drift parameter $\hat{d}$ is assumed to be a negative number. Hence, the expected value is decreasing over time, while the variance is increasing over time.

Now, by using the definition of the standard error estimate we obtain an estimation for σ. The definition is as follows[GK08]

Definition 3.4.1. The standard error estimate, denoted see_t is defined as

$$see_t = \sqrt{\frac{1}{T-2}\sum_{t=1}^{T-1}(\hat{k}_{t+1} - \hat{k}_t - \hat{\delta})^2}. \tag{3.15}$$

Where the index t is added to see_t to show its dependence on time.

By applying this estimator of see_t to equation 3.14 it can now be rewritten as

$$\begin{aligned} \hat{k}_t &= \hat{k}_T + \hat{\delta}t + \sqrt{t}\epsilon_t, \\ \epsilon &\sim \mathcal{N}(0, see_t^2) \end{aligned} \tag{3.16}$$

Now, by plugging the expression for $\hat{k}_t$ we obtained in equation 3.16 into equation 3.10, we obtain the forecast for log-mortality in the Lee-Carter model:

$$\log(\tilde{m}_{x,t}) = \hat{a}_x + \hat{b}_x(\hat{k}_T + \hat{\delta}t + \sqrt{t}\epsilon_t) = \hat{a}_x + U_{x,1}(\hat{k}_T + \hat{\delta}t + \sqrt{t}\epsilon_t), \tag{3.17}$$

where $\log(\tilde{m})$ is a $X \times T^{forc}$ matrix of forecasted log mortality rates, where X is the age limit and T^{forc} is the number of forecasted years. Since $\hat{a}_x$ and $\hat{b}_x$ remain constant in time and are calculated from empirical data, our expression depends only on the time-varying parameter, i.e. the matrix of forecasted mortality rates is a non-stationary time-series that is stochastic only in our time parameter t.

CHAPTER 4

Modeling and forecasting Norwegian mortality

4.1 Source of Data

All data used to calculate Norwegian mortality rates in this thesis is obtained from the Human Mortality Database, which is maintained by University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). The data is available from the website www.mortality.org or www.humanmortality.de (data downloaded in March 2022).

4.2 Description of Data

To perform our modeling and forecasting of Norwegian mortality, we require data for both death counts and exposure-to-risk. These data sets are stored respectively in files named NOR_deaths.txt and NOR_exposures.txt. Both files span the calendar years $1846 - 2020$ (one-year year grouping) and contains data for female, male and both genders combined, for one-year age groups $0, 1, 2, \ldots, 109, 110+$; where the last age group is all ages above 110. However, we will restrict the first part of our analyses to both genders combined, ages 0 to 90 for the years 1960 to 2020. The upper age limit of 90 years was selected in order to avoid the data sparseness at extreme ages and thereby reducing the uncertainty this entails. This limit is well above the life expectancy of both men and women in Norway in the year 2020, which is 81.5 and 84.9 years respectively.[1] Moreover, the projected Norwegian life expectancy in the year 2060 is 89.9 according to Statistics Norway[2], thereby making our limit suitable for comparison up until the year 2060.

Figure 4.1 below displays a three-dimensional plot of observed death counts for the Norwegian population for ages 0 to 90 over the years 1960 to 2020. The time age trend shows a sharp decline from year 0 (infant mortality), which

[1] Life expectancy in Norway. In: Public Health Report - Health status in Norway (online document). Oslo: Norwegian Institute of Public Health [updated 08.07.2021; read 03.01.2022]. Available from: https://www.fhi.no/nettpub/hin/samfunn/levealder/

[2] National population projections in Norway (online documents). Oslo: Statistics Norway (SSB) [updated 03.06.2020; read 03.02.2022]. Available from: https://www.ssb.no/en/befolkning/befolkningsframskrivinger/statistikk/nasjonale-befolkningsframskrivinger

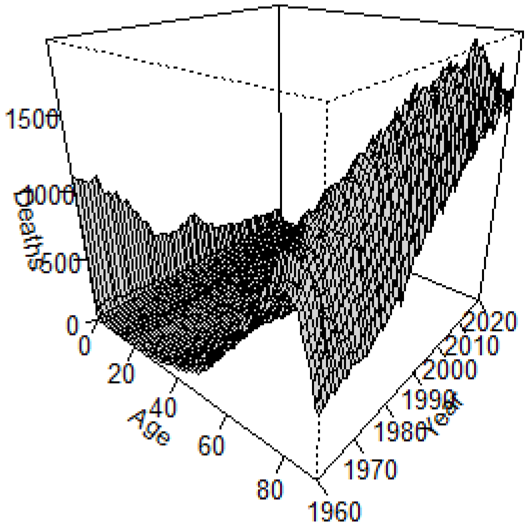

Figure 4.1: Observed number of deaths (1960-2020)

thereafter increases with higher ages, but dips around the age of 80 due to less and less people alive at ages above the life expectancy.

The figure 4.2 below shows the Norwegian exposure-to-risk for ages 0 to 90 and years 1960 to 2020. Predictably, the amount of individuals at risk decreases with time - less and less people are alive as age increases.

4.3 Mortality rates

We let $d_{x,t}$ and $E_{x,t}$ denote the empirical death count and the exposure-to-risk respectively, where $x = 0, \ldots, 90$ denotes the age period and $t = 1960, \ldots, 2020$ is the time period. From equation 2.9 we then have our crude mortality rate as an approximation of the crude hazard rate

$$\hat{r}_{x,t} = \hat{\mu}_{x+1/2,t+1/2} = \frac{d_{x,t}}{E_{x,t}},$$

where the notation is not to be confused with the one in the definition of the hazard rate in equation 2.5. Here, the index t represent the time period (year), and the additional $1/2$ tells us that this is the crude hazard rate for mid

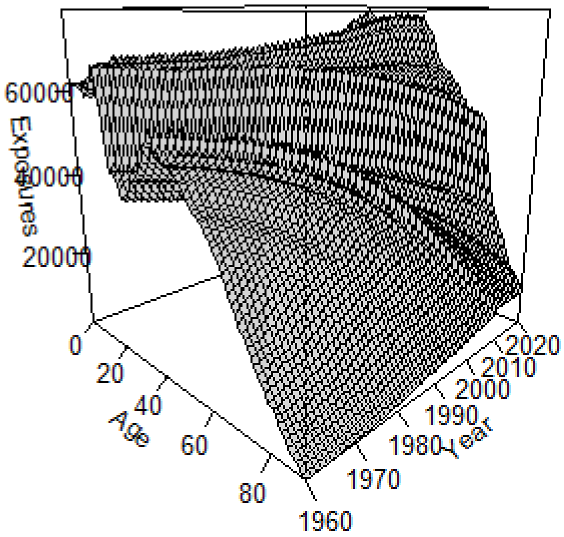

Figure 4.2: Observed exposure-to-risk (1960-2020)

year and mid age. The population of an area is non-stationary over the course of a year and the hazard rate fluctuates slightly. However by assuming the force of mortality remains constant for each age at a specific year over the course of the whole calendar year, we arrive at some useful approximations. Formalized, the assumption is as follows:

$$\mu_{x+s,t+v} = \mu_{x,t}, \quad \text{for all} \quad 0 \leq s, v < 1.$$

This is a often used assumption in the study of mortality which we will assume holds for our applications for our study of mortality in the Norwegian population.[Cai+09] The assumption implies some practical relationships between the force of mortality and mortality rates. We have the following:

1. $m_{x,t} = \mu_{x,t}$
2. $q_x = 1 - \exp(-\mu_{x,t}) = 1 - \exp(-m_{x,t})$,

where in the first relationship, $m_{x,t}$ is the central rate of mortality for an individual age x in year t and in the second relationship q_x is the yearly mortality rate in equation 2.3. Relationship 1 follows directly from equation

2.8 for constant $\mu_{x,t}$ and relationship 2 follows from equation 2.6 for constant $\mu_{x,t}$. By these assumptions our central mortality rate $m_{x,t}$ for age index x and year index t is therefore:

$$m_{x,t} = \hat{\mu}_{x,t} = \hat{\mu}_{x+1/2,t+1/2} = \frac{d_{x,t}}{E_{x,t}}.$$

Handling of zeros in the data set

In the next section, when fitting the Lee-Carter model, we will use the Singular Value Decomposition (SVD) on $\log(m_{x,t})$. Since Norway is a country with a relatively tiny population, there are zero values in our death counts data set. Hence, for some combinations of (x, t), we have $m_{x,t} = 0$. Since $\log(0)$ is undefined, we need a way to remedy this. For all (x, t) where this is the case, we approximate $m_{x,t}$ as such:

$$m_{x,t} = 0 \approx \frac{m_{x,t-1} + m_{x,t+1}}{2} = \frac{1}{2}\left(\frac{d_{x,t-1}}{E_{x,t-1}} + \frac{d_{x,t+1}}{E_{x,t+1}}\right).$$

I.e. we approximate with the mean of the previous and next years mortality for age x. Now, since our data set does not contain zero values in the first and last year (1960 and 2020 respectively) we do not need to handle the case of non-existent previous or preceding year. A flaw with this method is the scenario that the previous or preceding year had an extraordinary event that saw a huge spike in mortality rate for age x; though unlikely, given that the mortality rate at exact age x in year t is zero. For our data set, we do not have any such extreme events, thereby making this a valid method of approximating our zero values.

Figure 4.3 above shows a three-dimensional plot of log of observed mortality rates, where zero values are estimated as described above. For exact age x, the trend is decreasing mortality rates in the time-variable t. This trend is to be expected due to generally better standards of living, lower child mortality as a cause of better nutrition and medicinal advancements, as well as fewer work incidents, treatment of age-related medical conditions etc.

A Note on Covid-19

The Severe Acute Respiratory Syndrom coronavirus 2, SARS-CoV-2 or covid-19 for short, struck the world in 2020 and caused a worldwide epidemic. In Norway, the first case of SARS-CoV-2 was registered on the 26th of February 2020, and the first Norwegian death due to the disease was reported on the 12th of March the same year.[3] Since the year of outbreak is our last year of analysis, it is of interest to see how significant of an effect this epidemic had on the mortality rate. Deaths related to covid-19 in the Norwegian population (as well as worldwide) occurs much more frequently in the higher age groups: in year 2020, approximately 35% of the occurrences was registered in the 80 to 89 age group and approximately 28% registered in the 90+ age group. Furthermore, in

[3] Tjernshaugen, Andreas; Hiis, Halvard; Bernt, Jan Fridthjof; Braut, Geir Sverre; Bahus, Vegard Bø: koronapandemien i Store medisinske leksikon på snl.no. [updated 02.05.2022; read 02.05.2022]. Available form: http://sml.snl.no/koronapandemien

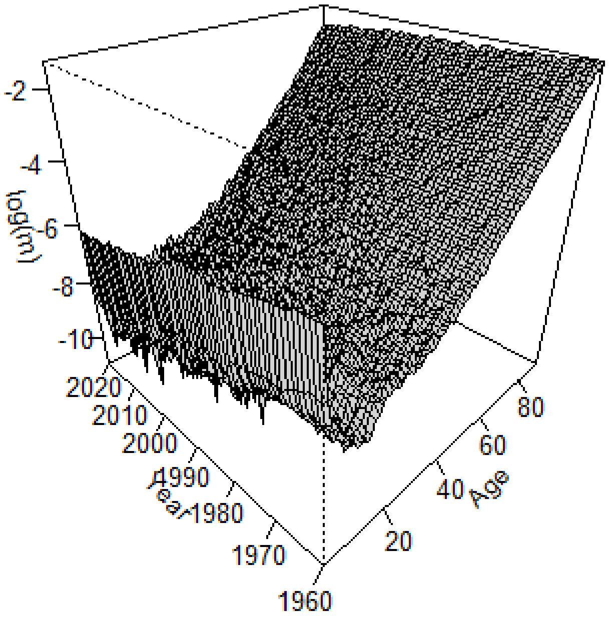

Figure 4.3: Log Mortality in Norwegian population (1960-2020)

2020 the Norwegian average age of dying related to a covid-19 infection (both genders together) was 81.5.[4] Therefore, to see the impact of the infection on the mortality rates, we will analyze the upper ages of our age interval.

Figure 4.4 shows the mortality curves for age 80, 85 and 90 plotted over years 1960 to 2020. This figure does not exhibit a huge spike in year 2020, the year covid-19 struck. However, we can see a slight increase in mortality from the previous year 2019. To see how significant the increase in mortality rates were, we can observe how the data points for 2020 relate to previous years.

Table 4.1 shows the Norwegian mortality rates for ages 80 to 90 in year 2020 and how they compare to the mortality rates for 2019, 2018 and 2017 given

[4] Sørlie Strøm, Marianne; Raknes, Guttorm: Tall for covid-19 assosierte dødsfall i Dødsårsaksregisteret i 2020. [published 10.06.2021; read 28.04.2022]. Available from: https://www.fhi.no/hn/helseregistre-og-registre/dodsarsaksregisteret/tall-for-covid-19-assosierte-dodsfall-i-dodsarsaksregisteret-i-2020/

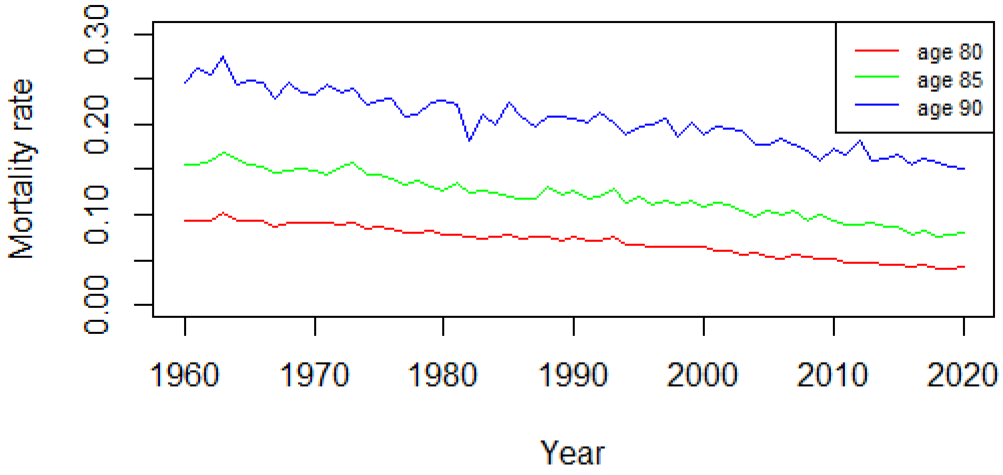

Figure 4.4: Mortality rates for high ages (1960-2020)

Age x	Mortality	1 year difference	2 year difference	3 year difference
80	0.04163054	4.70 %	1.36 %	-6.61 %
81	0.04314581	-7.40 %	-12.86 %	-19.19 %
82	0.05290193	2.50 %	-2.80 %	-0.87 %
83	0.05942208	-8.43 %	-8.74 %	-0.97 %
84	0.06966374	-1.50 %	-3.50 %	-5.23 %
85	0.07977311	2.46 %	6.08 %	-2.77 %
86	0.08281899	-6.18 %	-10.63 %	-13.53 %
87	0.09796518	-5.83 %	-4.79 %	-9.54 %
88	0.11546632	-2.40 %	-4.57 %	-2.14 %
89	0.13369501	6.00 %	-7.06 %	-7.36 %
90	0.15011664	-1.71 %	-4.24 %	-7.66 %

Table 4.1: Mortality rate for ages 80 to 90 in year 2020 and how they compare to the three previous years.

in percentage increase or decrease (rounded to two decimals points). Observe that for these high ages, the mortality rates decreased overall compared to the previous year 2019; although age 80 and 89 show a not so insignificant increase of 4.70% and 6.00% respectively. The mortality rates for 2020 show a larger percentage decrease compared to the 2018 levels, and a even larger compared to the 2017 levels. This reflects the trend of decreasing mortality observed in figure 4.4.

From this we can conclude that the mortality rates for higher ages in 2020 has not seen a significant increase from previous years, but rather follows the decreasing trend in mortality over the years. However, since we are dealing with all-cause mortality, we cannot draw any clear conclusions of how large of an effect covid-19 had on the total mortality: there might be a decline in other causes of mortality off-setting the effect of covid-19 on the mortality rates. Since the great part of covid-19 related deaths occurred in the higher age range, above the life expectancy for the Norwegian population, which is an age group with a statistical high probability of dying that same year, a lot of the cases might be a result of comorbidities which would with a high probability result in death in 2020, with or without the disease. Further studies can be made into the even higher ages of 90+, which would be expected to be even more prone to dying from covid-19. It also remains to see what impact the infection had on mortality rates for 2021 and 2022 by the end of the year.

4.4 Fitting the Lee-Carter model

In this section we fit the Lee-Carter model on the empirical mortality rates as explained in section 3.2. All calculations pertaining to this procedure will be done using the programming language `R`[R C22]. First we calculate the general pattern of mortality a_x by simply taking `rowMeans(log(crude.mort))` to get the average log mortality at every age over years t, where `crude.mort` is our (91×61)-matrix of mortality rates. Then, by using `R`'s built in `svd()`-function on matrix `log(crude.mort)-ax` we get our U, V along with the singular values σ_i: The function returns a (91×91) `u`-matrix, a (61×61) `v`-matrix and a 61 length vector d corresponding to the singular values. From these we extract `u[,1]`, `v[,1]` and the principal component `d[1]` to calculate `bx = -u[,1]` and `kt = -d[1]*v[,1]`. The minus sign is added since the `svd()`-function in `R` results in $\hat{k}_t$ ordered from lowest to highest value. By putting a minus in front we reverse this effect, which we can do because of the invariancy of the parameters explained in section 3.2, equation 3.4. Thereby we gather the Lee-Carter approximated log mortality rates in a matrix where $\log(m_{x,t}) = a_x + b_x k_t$ as in equation 3.3.

Table 4.2 shows the estimated Lee-Carter $\hat{a}_x$ and $\hat{b}_x$ values for the whole age span 0 to 90.

Table 4.3 shows the estimated Lee-Carter $\hat{k}_t$ for the whole time span 1960 to 2020.

Figure 4.5 shows a surface plot of the Lee-carter fitted mortality rates for years 1960 to 2020 and ages 0 to 90. The trend is here the same as the observed mortality rates observed in figure 4.3, but the curve is predictably smoother.

age	ax	bx
0	-5.134506	0.188023
1	-7.392926	0.220653
2	-7.985152	0.218570
3	-8.184486	0.233794
4	-8.398132	0.219707
5	-8.467590	0.205210
6	-8.574621	0.229066
7	-8.634246	0.199080
8	-8.782404	0.216085
9	-8.885317	0.177285
10	-8.884986	0.169512
11	-8.848040	0.156576
12	-8.773475	0.139314
13	-8.606419	0.136290
14	-8.443385	0.143438
15	-8.178761	0.139431
16	-7.826874	0.115324
17	-7.651980	0.113840
18	-7.369245	0.083741
19	-7.348706	0.082835
20	-7.331423	0.058862
21	-7.409303	0.070690
22	-7.358805	0.054308
23	-7.354796	0.063086
24	-7.377486	0.051510
25	-7.371457	0.047104
26	-7.356222	0.043422
27	-7.326044	0.044192
28	-7.319637	0.050232
29	-7.272468	0.047078
30	-7.285358	0.062589
31	-7.208253	0.053200
32	-7.159042	0.055150
33	-7.123876	0.064690
34	-7.057406	0.062149
35	-7.000200	0.068092
36	-6.926974	0.071552
37	-6.857773	0.076362
38	-6.792036	0.075649
39	-6.713966	0.078767
40	-6.634209	0.079237
41	-6.565089	0.075244
42	-6.478845	0.084280
43	-6.368383	0.084542
44	-6.283931	0.092489
45	-6.169564	0.090070
46	-6.086336	0.084258
47	-5.989959	0.082899
48	-5.895039	0.086326
49	-5.785095	0.084254
50	-5.692399	0.083017
51	-5.590504	0.080668
52	-5.506006	0.081724
53	-5.412823	0.081037
54	-5.306463	0.080376
55	-5.200500	0.080956
56	-5.128351	0.083349
57	-5.023178	0.079858
58	-4.936686	0.085332
59	-4.828980	0.079598
60	-4.736759	0.079529
61	-4.635613	0.082154
62	-4.533971	0.080241
63	-4.439240	0.081692
64	-4.338225	0.082717
65	-4.253655	0.082807
66	-4.148528	0.079664
67	-4.055141	0.080362
68	-3.950679	0.080773
69	-3.858538	0.081196
70	-3.758309	0.086104
71	-3.643393	0.083076
72	-3.549114	0.081791
73	-3.443930	0.082763
74	-3.342258	0.082879
75	-3.228993	0.079764
76	-3.130571	0.081658
77	-3.019355	0.078202
78	-2.906721	0.076493
79	-2.795599	0.074579
80	-2.687495	0.073077
81	-2.574719	0.069390
82	-2.462630	0.067731
83	-2.351386	0.064229
84	-2.241218	0.062388
85	-2.135019	0.059077
86	-2.030803	0.055752
87	-1.919134	0.052318
88	-1.812568	0.049439
89	-1.702510	0.045609
90	-1.600068	0.040650

Table 4.2: Lee-Carter a_x and b_x values

year	kt
1960	4.635218
1961	4.492249
1962	4.524585
1963	4.649986
1964	4.370957
1965	4.449488
1966	4.026584
1967	4.013003
1968	4.028702
1969	4.403771
1970	3.917311
1971	4.062871
1972	3.824077
1973	3.704487
1974	3.629101
1975	3.461861
1976	3.010225
1977	2.714977
1978	2.598229
1979	2.452624
1980	1.866983
1981	1.793721
1982	1.771269
1983	1.685691
1984	0.973362
1985	1.722648
1986	1.455857
1987	1.102948
1988	1.062193
1989	0.680112
1990	1.127103
1991	0.228657
1992	0.191719
1993	-0.042112
1994	-0.575235
1995	-0.647838
1996	-1.210817
1997	-0.678415
1998	-1.378528
1999	-1.010720
2000	-1.175970
2001	-1.981414
2002	-1.244238
2003	-1.745171
2004	-2.600832
2005	-2.577627
2006	-3.302878
2007	-4.057676
2008	-3.577334
2009	-3.734439
2010	-4.314564
2011	-3.879303
2012	-4.701559
2013	-5.490387
2014	-5.375768
2015	-6.157458
2016	-6.230144
2017	-5.841282
2018	-6.546524
2019	-5.961992
2020	-6.592342

Table 4.3: Lee-Carter k_t values

Figure 4.6 displays plots of the estimated parameters of the Lee-Carter model, $\hat{a}_x$, $\hat{b}_x$ and $\hat{k}_t$ plotted against age, age (x) and years (t) respectively. As expected, the general pattern of mortality $\hat{a}_x$ increases with age. From table 4.2 we see that $\hat{a}_x$ decreases from about -5.13 at year 0 to about -8.89 at year 9, from whence it then increases to -16.00 for the 90 year group.

The estimated age specific parameter $\hat{b}_x$ decreases rapidly from the early ages and then flattens out after year 20. High values of $\hat{b}_x$ means that mortality varies significantly for a change in time index variable $\hat{k}_t$. The opposite, low values of $\hat{b}_x$ means that mortality varies less with changing time variable $\hat{k}_t$. The high values at young ages can be interpreted as the mortality at these ages have seen a significant change over the years. It is well known that child mortality and youth mortality have decreased significantly the last 60 years.

The values for the time-varying parameter $\hat{k}_t$ are plotted at the top left

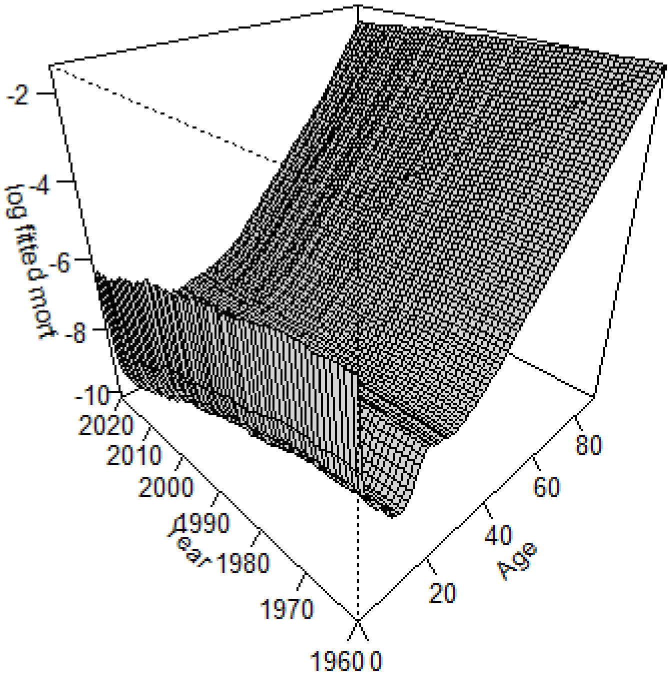

Figure 4.5: LC fitted mortality rates (1960-2020)

of figure 4.6. These values capture the main time trend of mortality in the Norwegian population over the years 1960 to 2020. As expected, the trend is decreasing, but the values fluctuate slightly from year to year as shown in table 4.3.

The bottom right plot of figure 4.6 shows the Lee-Carter approximation together with the log of empirical mortality rates for age 50 over the year span 1960 to 2020. The Lee-Carter shows a good fit against the data point and has the same downward trends as the empirical $\log(m_{x,t})$. It is however not able to capture some of the early years with great accuracy. This is not necessarily a flaw in the model, since over-fitting can lead to worse prediction capabilities.

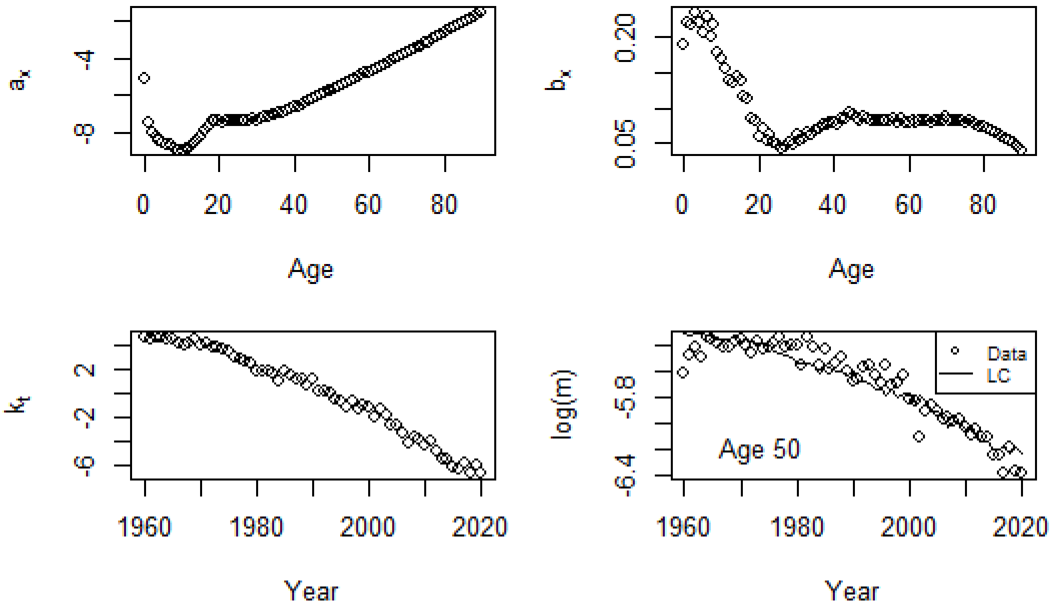

Figure 4.6: Data and parameter estimates in the Lee-Carter model for the Norwegian population. Data for ages 0 to 90 and years 1960 to 2020.

Goodness of Fit

From the preliminary analysis into Lee-Carter fitted on Norwegian mortality data, it seemed to fit the data points with a fairly good accuracy, capturing both the time and age trends in the data. An important question is how good this fit really is. To assess the *goodness of fit* of the Lee-Carter model, we employ a well known statistical measurement.

Variance of Distance Explained

To measure the goodness of fit to empirical data, Lee and Carter (1992) used the *ratio of variance explained*[LC92]. By first calculating the variance of the difference between empirical mortality $m_{x,t}$ and the fitted mortality $\hat{m}_{x,t}$ and then dividing this by the variance of the empirical data, the result is a ratio of *badness of fit*. To obtain the goodness of fit, the ratio is therefore subtracted from 1. For age x, it is defined η_x and is explicitly given as follows:

$$\eta_x^2 = 1 - \frac{V\left[||m_{x,t} - \hat{m}_{x,t}||\right]}{V\left[||m_{x,t}||\right]} = 1 - \frac{\sum_{t=1}^{T}(m_{x,t} - \hat{m}_{x,t})^2}{\sum_{t=1}^{T}(m_{x,t} - \bar{m}_x)}, \tag{4.1}$$

in which $\bar{m}_x$ denotes the mean value of the empirical mortality, $\exp(a_x)$. Observe that this expression is the same as the *coefficient of determination* R^2 in regression analysis. The resulting η_x^2 is a percentage proportion $0 < \eta_x^2 < 1$ of the models ability to explain the variance of the actual, empirical data.

The ratio of variance explained η_x^2 will be calculated for the Lee-Carter estimate of Norwegian mortality. For convenience, ages x will be grouped together in 9 groups: $0-10$, $11-20$, $21-30$, $31-40$, $41-50$, $51-60$, $61-70$, $71-80$ and $81-90$. To estimate η^2 for these groups, we take the mean of the η_x^2 for each age x included in each chosen grouping.

Age group	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90
η^2	0.892	0.699	0.497	0.733	0.868	0.910	0.946	0.971	0.955

Table 4.4: Ratio of variance explained in LC by age groups

The resulting η^2 values for the mentioned aged groups are listed in table 4.4 (rounded to three decimal points). From age 51 the Lee-Carter model explains a lot of the variance of observed data, with $\eta^2 > 90\%$ for these age groups. It also captures the variance for the child mortality group, that is ages 0-10 with great accuracy, explaining 89% of the variance. It does, however, fail to sufficiently capture the variance for age groups 11-20 and 21-30, with a ratio of variance of 70% and 50% respectively. By inspecting our data for η_x we see that our lowest η_x^2 value is at age 24, only accounting for 36% of the total variance, while our highest η_x^2 value is at age 80, where LC accounts for 98% of the total variance. The huge difference in explanation capability is visualized in figure 4.7 and figure 4.8, which shows the empirical log mortality together with the Lee-Carter approximation for age 24 and 80 respectively. There has evidently been a lot of variation in Norwegian mortality for age 24 over the years, making it hard to capture in the standard Lee-Carter model. In comparison, mortality for age 80 follows a almost linearly decreasing time trend with very little variation, making it easier to capture for the model.

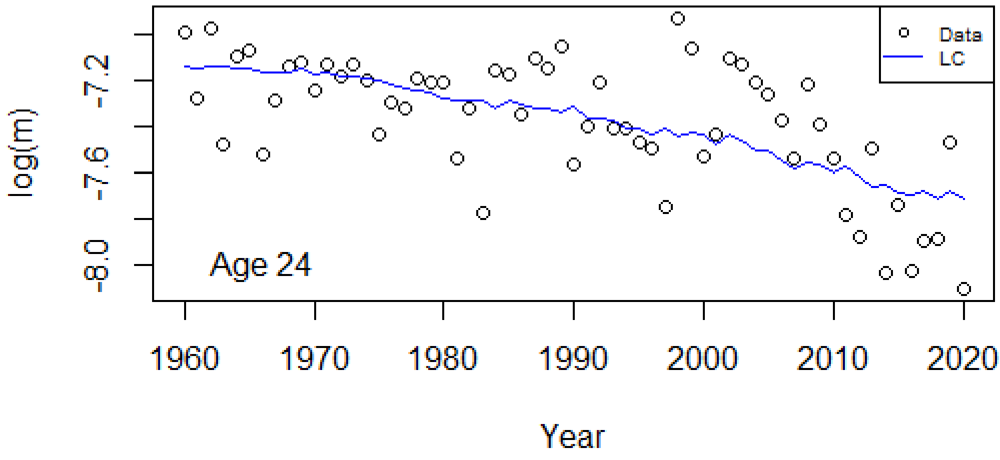

Figure 4.7: Lee-Carter and empirical data for age 24, years 1960-2020.

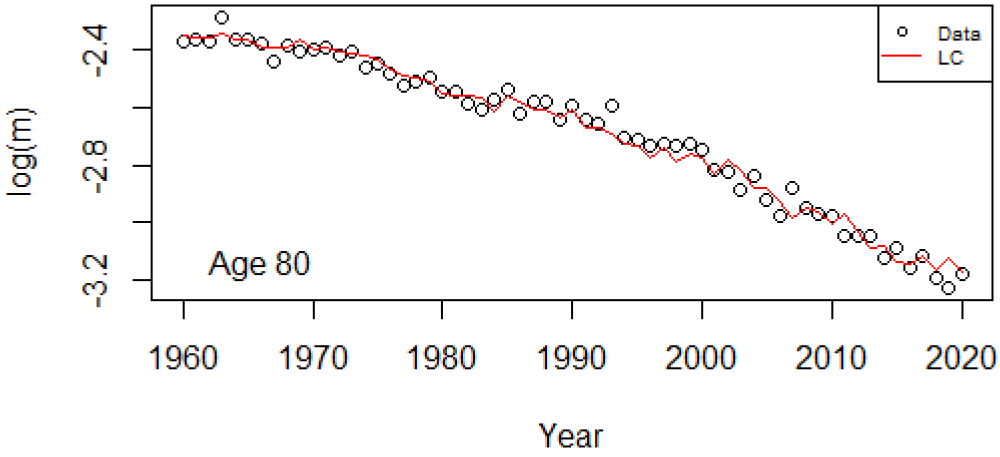

Figure 4.8: Lee-Carter and empirical data for age 80, years 1960-2020.

4.5 Forecasting Norwegian Mortality

In this section we forecast the Norwegian mortality using the Lee-Carter method as described in section 3.4. We are interested in forecasting Norwegian mortality 40 years ahead from our empirical data set, that is for years 2021 to 2060. This is done by forecasting the time-varying parameter $\hat{k}_t$ for $t = 2021 \ldots 2060$. In R this is done by first calculating $\hat{\delta}$ from equation 3.13 and the standard error estimate see_t from equation 3.16 as such:

```
# Calculating standard error
d.hat <- (kt[T] - kt[1])/(T - 1)
se.sum <- rep(0, length = T-1)
for (i in 1:T-1){
        se.sum[i] <- (kt[i+1] - kt[i] - d.hat)^2
}
se.hat <- sqrt((1/(T - 2))*sum(se.sum))
```

where the integer T is the final time point of the empirical Norwegian mortality data $t = 0, \ldots, 61$ (i.e.year 2020) and se.hat is the standard error estimate see_t. Thereby, following the Lee-Carter algorithm, the forecasted $\hat{k}_t$ are calculated with the snippet below:

```
# Forecast future kt
set.seed(seedn)
kt.hat <- matrix(nrow = forc.t, ncol = 1)
for (i in 1:forc.t){
kt.hat[i] <- kt[T] + i*d.hat
error_term <- rnorm(1, 0, se.hat^2)
kt.hat[i] <- kt.hat[i] + sqrt(i)*error_term
 }
```

Here, `error_term` is the ϵ_t in equation 3.16, which is assumed normal distributed with expectation 0 and variance see_t^2. This distribution is easily simulated in `R` with the function `1, 0, se.hat`$\hat{2}$. The parameter `forc.t` denotes the number of years forecasted into the future, which here is 40; while the parameter `T` equals the time period length of the estimated values, 61. For replicability we set the seed with the function `set.seed(seedn)`, where `seedn = 760` in this case. Figure 4.9 shows the forecasted $\hat{k}_t$ values for years 2021 to 2060. We see the same downward trend in the forecasted values as in the plot for fitted $\hat{k}_t$ in figure 4.6. This downward time trend becomes even more apparent in figure 4.10 which shows a line-plot of both fitted and forecasted $\hat{k}_t$ values in green and red respectively: the line through the forecasted values follow the same decrease where the fitted values end in year 2020, though it exhibits more variation due to the uncertainty of the random ϵ_t.

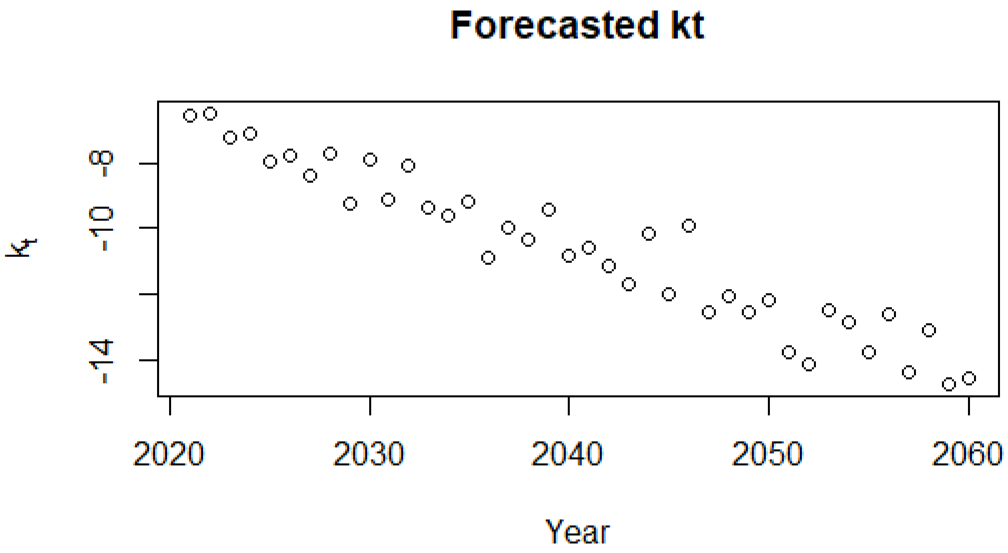

Figure 4.9: Lee-Carter forecasted $\hat{k}_t$ for years 2021 to 2060.

Now that we have calculated the forecasted $\hat{k}_t$ values, it remains only to plug these values into Lee-Carter equation (3.17) to obtain our matrix of forecasted log mortality rates given by

$$\log(\tilde{m}_{x,t}) = \hat{a}_x + \hat{b}_x\hat{k}_t, \quad t = 62\ldots91,$$

where $\tilde{m}_{x,t}$ denotes the forecasted mortality rates at age $x = 0\ldots90$ and $t = 62\ldots92$ represents the forecasted year integers for 2021 to 2060. The resulting surface of log mortality is plotted in figure 4.11. This plot has the same general time and age trends as the surface plot of the Lee-Carter approximation on empirical data in figure 4.5. There are, however, some differences in the shapes of these two curves: While the approximated log mortality exhibits an increase up until year 20, this is much more pronounced in the forecasted values at the same age, which almost reach the mortality levels at the very first

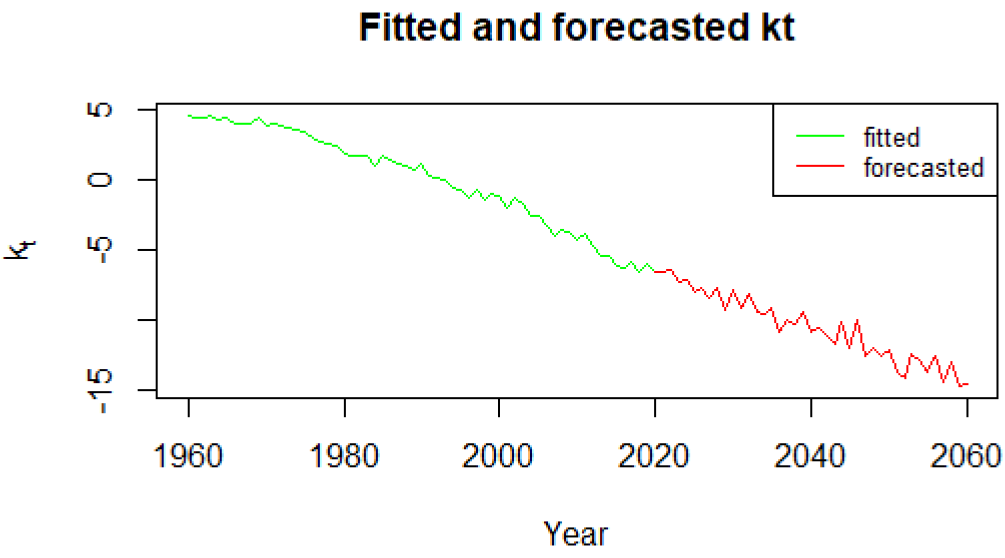

Figure 4.10: Lee-Carter fitted and forecasted $\hat{k}_t$

ages of analysis. However, the forecasted early mortality rates are significantly lower than the fitted ones; as well as the overall mortality for age and year being lower. The visualization of how the log mortality rates develop from the estimated values to the span of the forecasted years is shown in figure 4.12. In this figure, the curve of Lee-Carter estimated log mortality for ages 0 to 90 in year 2020 is compared to the corresponding values for forecasted years 2030, 2040, 2050 and 2060. For all ages the mortality curves are decreasing in time, but the largest decrease is seen in the early ages, with the greatest difference in log mortality between 2060 and 2020 being -1.861521 at age 3. This forecasted decrease in child mortality reflects both worldwide empirical studies on mortality over the years (see e.g. Ahmad et al.[ALI00]), but also the empirical Norwegian crude mortality rates from our data sets.

A point of interest from an actuarial perspective, is to see how the mortality rate for a certain age x is forecasted to evolve in the future. Revisiting the case of a 50 year old individual from the bottom right panel of figure 4.6, we plot the empirical mortality rates together with fitted and forecasted estimates. The result are shown in figure 4.13.

Comparison to K2013

The $8th$ of march 2013 the financial supervisory authority of Norway (Finanstilsynet) published a new basis for estimating mortality in age- and survivor pensions in collective pension insurance.[5] The document, called K2013

[5] New mortality basis in collective pension insurance (K2013) (online letter). The financial supervisory authority of Norway [published 08.03.2013; updated 27.03.2019; read 19.03.2022].

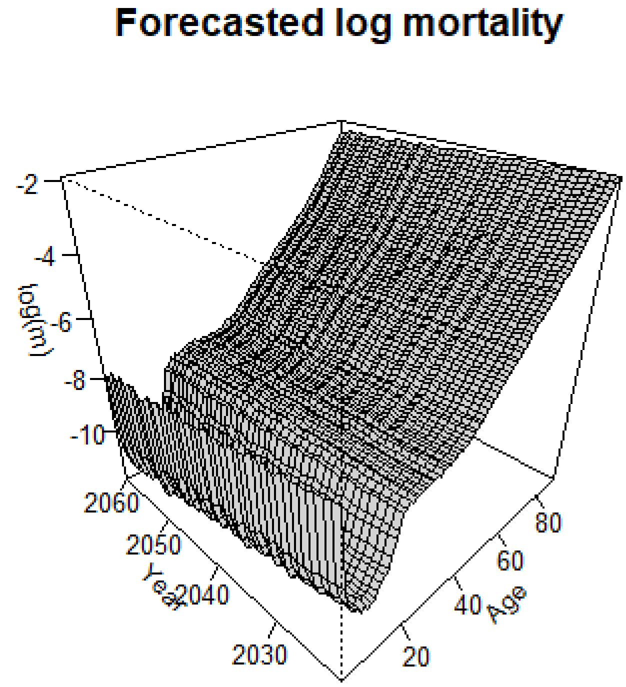

Figure 4.11: Forecasted log mortality for years 2021 to 2060

for short, proposes the following mortality model for age x in calendar year t:

$$\mu_{Kol}(x,t) = \mu_{Kol}(x,2013) * \left(1 + \frac{w(x)}{100}\right)^{t-2013} \quad , \quad t \geq 2013 \qquad (4.2)$$

where $\mu_{Kol}(x,2013)$ is the mortality for an insured of age x in year 2013. Furthermore, $w(x)$ denotes a decline in mortality function, given for the different genders as such

$$w(x) = min(2.671548 - 0.172480x + 0.001485x^2, 0), \quad \text{for men}$$
$$w(x) = min(1.287968 - 0.101090x + 0.000814x^2, 0), \quad \text{for women}$$

In this subsection a comparison between $\mu_{Kol}(x,t)$ and the forecasted Lee-Carter mortality rates we obtained will be made for year $t = 2021, \ldots, 2060$

Available from: https://www.finanstilsynet.no/nyhetsarkiv/pressemeldinger/2013/nytt-dodelighetsgrunnlag-i-kollektiv-pensjonsforsikring/

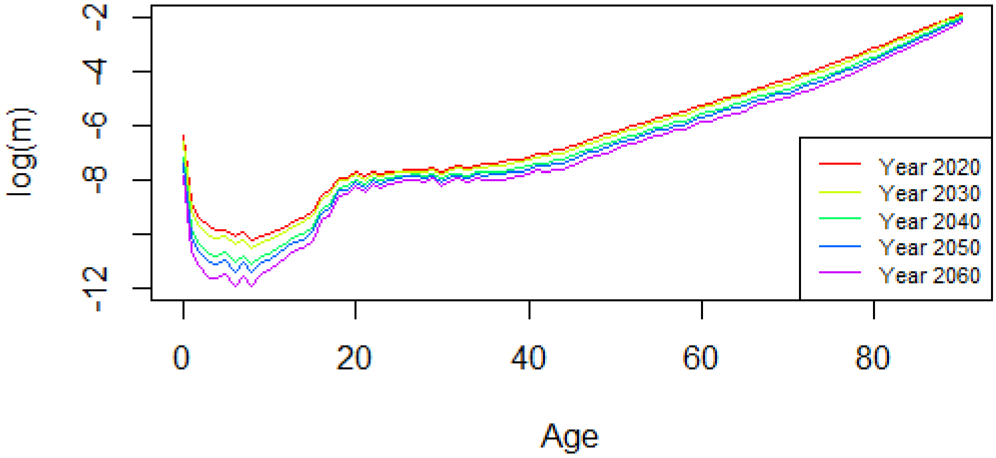

Figure 4.12: log mortality for 2020, 2030, 2040, 2050 and 2060

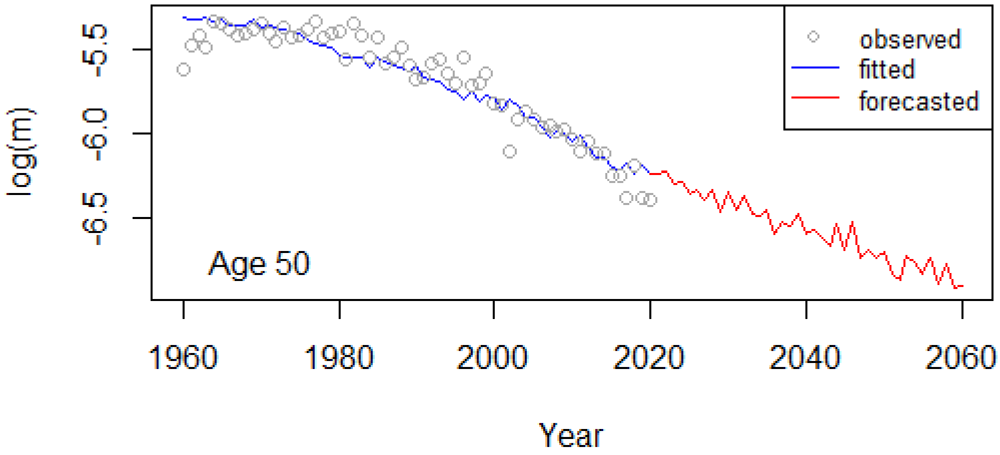

Figure 4.13: Observed, estimated and forecasted log mortality for age 50

will be made. Since the forecasted mortality rates are both genders combined, for each age x we will take the mean of $w(x)$ for both genders as such:

$$\hat{w}_x = min(mean(2.671548 - 0.172480x + 0.001485x^2, 1.287968 - 0.101090x + 0.000814x^2), 0)$$

Figure 4.14 shows the Lee-Carter forecasted $\log(\tilde{m}_{x,t})$ mortality rates compared to $\log(\mu_{Kol}(x,t))$ for ages 10, 25 and 75 for years $2021 - 2060$. The seemingly discrepancy between forecasted Lee-Carter and K2013 rates for young ages is due to the logarithmic measure of mortality. Figure 4.15 plots estimated q_x from both of the models forecasts for year 2060 and ages $0 - 90$. From this plot we see that the models track each other very nicely at young ages. The higher the ages though, the Lee-Carter model predicts increasingly higher q_x than what K2013 does. Using the K2013 model of mortality forecasting in calculation of pension insurance policy will therefore yield more conservative reserve estimates and premiums than the Lee-Carter will, since it assumes a lower probability that the policy holder will die during future years than the Lee-Carter does, thereby increasing the present value of the promised cash flows to the insured. The opposite is true for a life insurance policy that promises a certain amount if the insured dies.

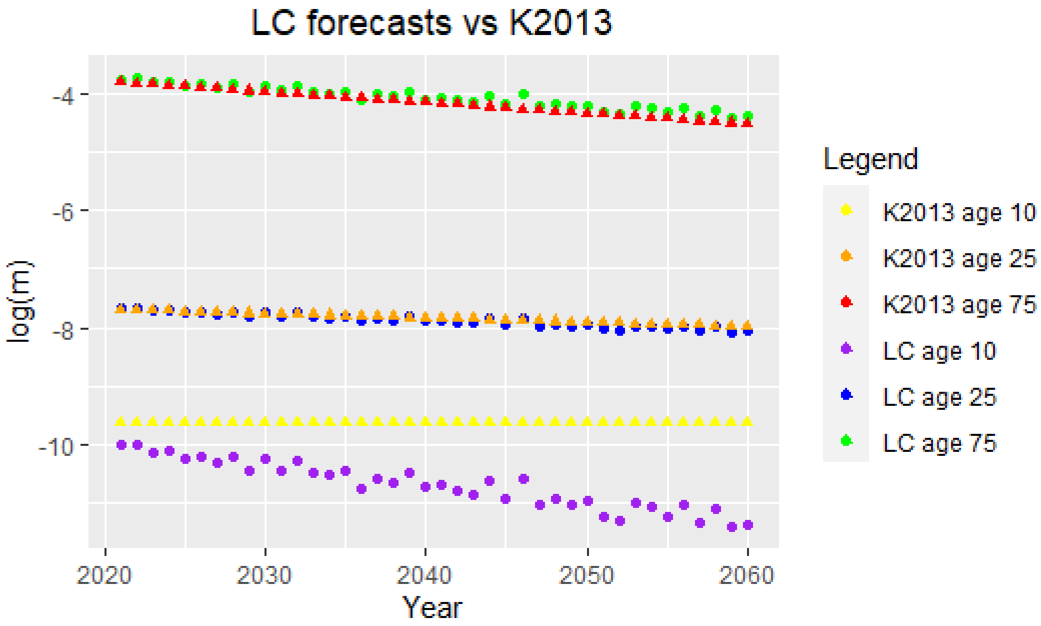

Figure 4.14: LC and K2013 forecasted log mortality for ages 10, 25, 75.

4.6 Forecast Accuracy

In the last section we forecasted mortality rates up until year 2060. The question we now ask ourselves is how good the Lee-Carter model is at predicting future

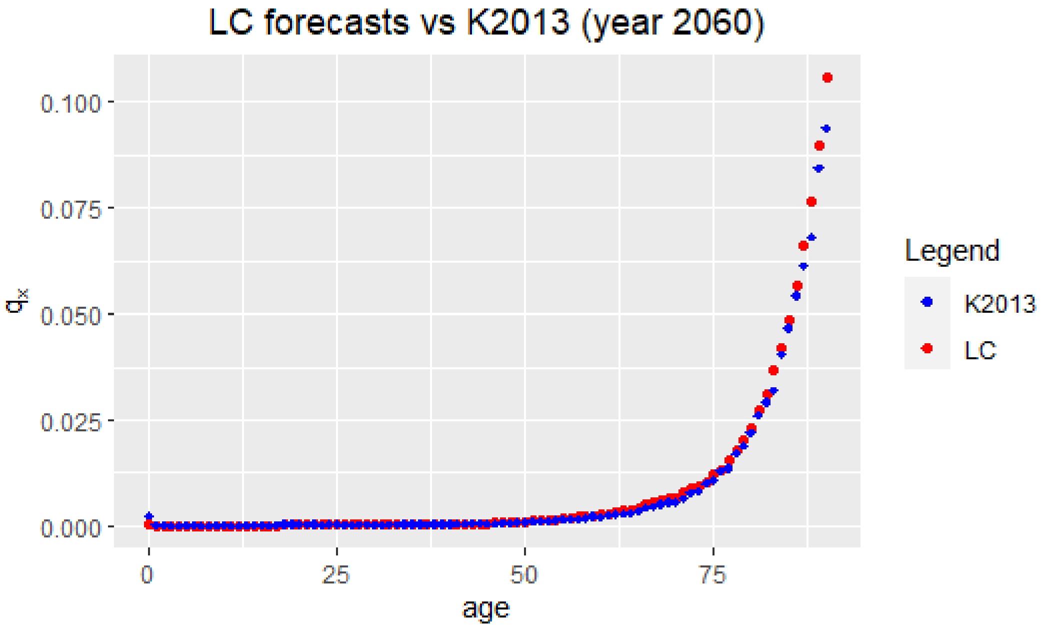

Figure 4.15: LC and K2013 forecasted q_x for year 2060.

mortality rates for Norwegian data. To find out, we partition our data files NOR_deaths.txt and NOR_exposures.txt into two sets, one training set and one test set:

- **Training set**: On this set we fit the Lee-Carter model. For the time period we chose the years $1920 - 1990$. The lower end of the interval, 1920, is chosen partly to have enough data to fit the Lee-Carter model, but also to avoid the Spanish flu, which inflicted a lot of casualties in the Norwegian population during the years 1918 and 1919. After we obtain our Lee-Carter estimates, we forecast the mortality rates for the time-period of the test set.

- **Test set**: This test set will consist of data for years $1991 - 2020$, the remainder of our data sets for the Norwegian population. It will be used to assess the accuracy of Lee-Carter forecasts by comparing the values to the forecasts made based on the training set.

Both the training set and test set will include ages 0 to 90. By the two partitions of our exposures and deaths files, we calculate the mortality $m_{x,t} = (D_{x,t}/E_{x,t})$ as before. The code for the partitioning and mortality calculation is given below:

```
# Training set from year 1920-1990
D.train <- Deaths[Age <= 90, (Year >= 1920) & (Year <= 1990)]
E.train <- Expos[Age <= 90, (Year >= 1920) & (Year <= 1990)]
mort.train <- get.Crude.M(D.train, E.train)

# Test set from year 1991-2020
D.test <- Deaths[Age <= 90, (Year > 1990) & (Year <= 2020)]
E.test <- Expos[Age <= 90, (Year > 1990) & (Year <= 2020)]
mort.test <- get.Crude.M(D.test, E.test)
```

The function `get.Crude.M(d, e)` in the snippet above returns the mortality rates $m_{x,t}$, with the approximations for zero values as described in section 4.3. Using the procedure for Lee-Carter estimations on the training set, we obtain our fitted $\log(\hat{m})_{x,t}$ values for ages $0, \ldots, 90$ and years $1920, \ldots, 1990$. Figure 4.16 shows the resulting surface mortality plot. Observe that the plot shows an increase for almost all ages around the years of the second World War (1940-1945). This increase is made more apparent in figure 4.17, where the estimated Lee-Carter parameters $\hat{a}_x$, $\hat{b}_x$ and $\hat{k}_t$ are plotted. It shows a significant uptick in the time-varying parameter $\hat{k}_t$ for these years. The plot also displays the empirical and estimated mortality for age 50 over the duration of the time span $1920 - 1990$, showing the effect of $\hat{k}_t$ on estimated mortality for this age during the war years.

Figure 4.18 shows the bottom right panel of figure 4.17 by itself, inspecting the mortality for age 50 in our training set more closely. The Lee-Carter estimation seems to fit the data quite good, but underestimates the mortality for the first and last years of our year interval. There are some outliers in the mortality set, with the most apparent here being for year 1960, which is very low compared to the trend. By comparing it to the plot we obtained for the time span $1960 - 2020$ in the lower right panel of figure 4.6, we can conclude that for age 50 the training set contains more variation for age 50. From the surface mortality plot of Lee-Carter fitted values for years $1920 - 1990$ we expect this to be the case for other ages also, compared to $1960 - 2020$.

Now that we have our Lee-Carter estimated $\log(\hat{m}_{x,t})$ on our training set, we follow the forecasting procedure from section 4.5 and obtain our forecasted $\hat{k}_t$ and mortality $\log(\tilde{m}_{x,t})$ for years $1991 - 2020$. The resulting forecasted $\hat{k}_t$ are plotted with the estimated ones in figure 4.19.

Confidence Intervals

To measure the credibility of future predictions, we apply confidence intervals to our forecasted log mortality rates. This is done by multiplying each forecasted data point $\log(\tilde{m}_{x,t})$ with a factor given as[LC92]

$$\log(\tilde{m}_{x,t}) \times \exp(-1.96\hat{b}_x see_t) \leq \log(\tilde{m}_{x,t}) \leq \log(\tilde{m}_{x,t}) \times \exp(1.96\hat{b}_x see_t).$$

Figures 4.20, 4.21 and 4.22 show the forecasted log mortality rates with confidence intervals for age $0 - 90$ in year 2000, 2010 and 2020 respectively, together with the corresponding empirical log rates from the test set. For all

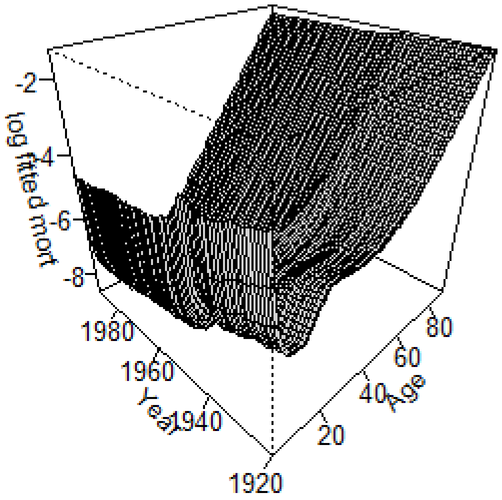

Figure 4.16: Surface plot LC $\log(\hat{m}_{x,t})$ on training set. Data: Norwegian mortality for ages $0-90$ and years $1920-1990$.

three years, the empirical values lie inside the confidence intervals up until about age 50, thus the Lee-Carter model is able to accurately predict the mortality rate for these ages. From age 50 onward the Lee-Carter predictions overestimate the mortality rate, and almost all the data points are outside of the confidence interval for the predictions. A factor contributing to worse predicting capabilities might be the higher mortality rate during the war. Another possibility might be that these years have seen a faster decrease in mortality for high ages than what the training set's trend *could* imply.

Measurement of Forecast Errors

We define by $e_{x,t}$ the forecast errors - that is the difference between the observed mortality and the predicted mortality - as such

$$e_{x,t} = m_{x,t} - \tilde{m}_{x,t}, \tag{4.3}$$

where for our case, $x = 0, \dots, 90$ denotes the age and $t = 1991, \dots 2020$ denotes the forecasted years. From our definition we see that positive $e_{x,t}$ implies $m_{x,t} > \tilde{m}_{x,t}$, negative $e_{x,t}$ implies $m_{x,t} < \tilde{m}_{x,t}$ and zero valued $e_{x,t}$

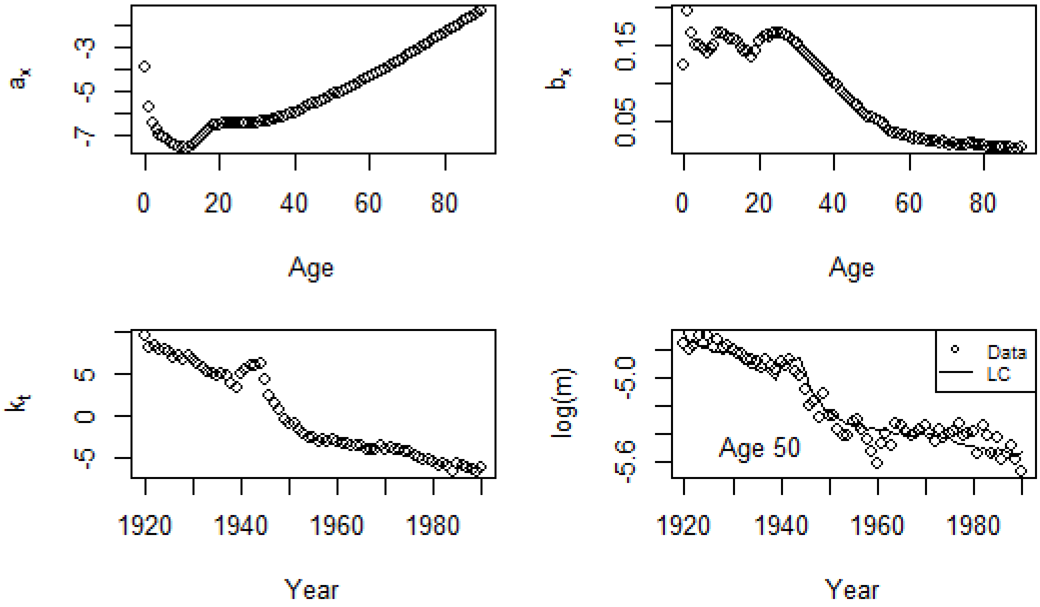

Figure 4.17: LC parameters and estimated mortality for age 50, years 1920 – 1990.

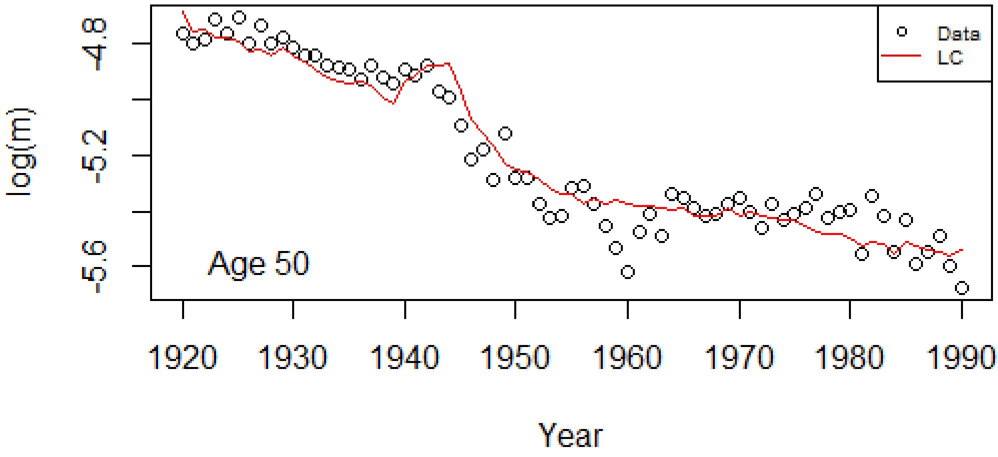

Figure 4.18: LC estimated mortality for age 50, years 1920 – 1990.

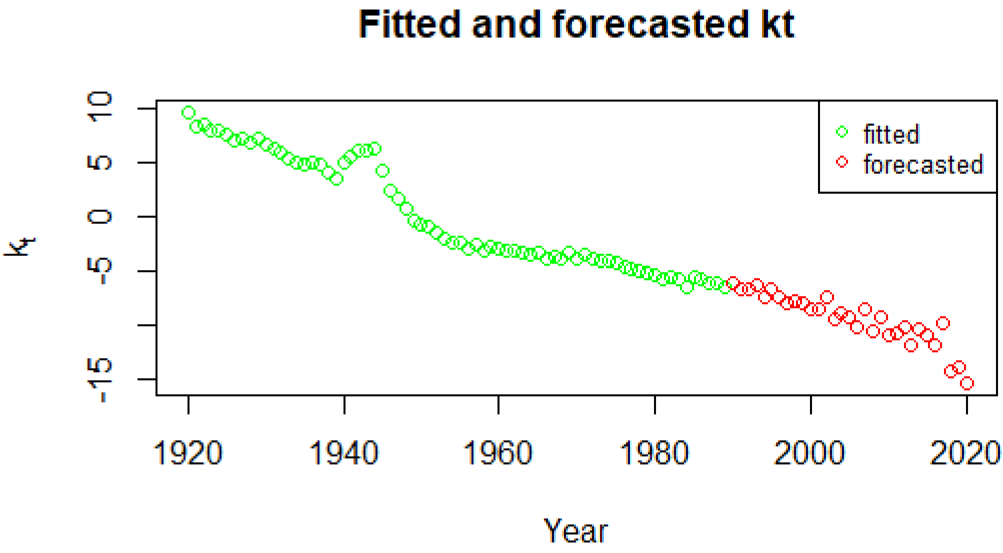

Figure 4.19: Estimated and forecasted $\hat{k}_t$ for 1920 – 2020.

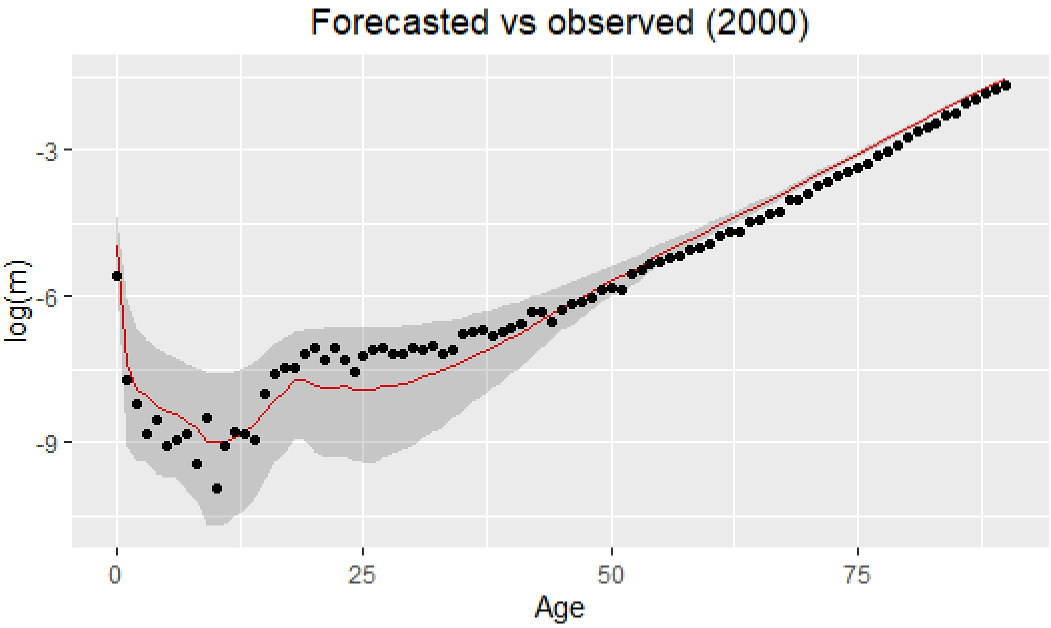

Figure 4.20: LC forecasted and empirical $\log(m)$ for 2000, with 95% CI.

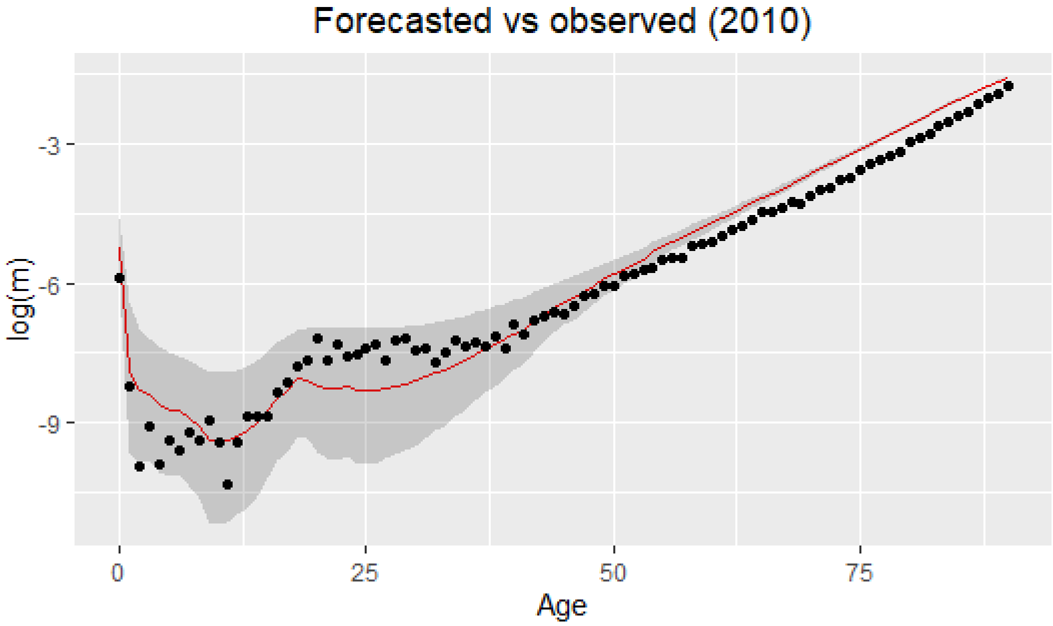

Figure 4.21: LC forecasted and empirical $\log(m)$ for 2010, with 95% CI.

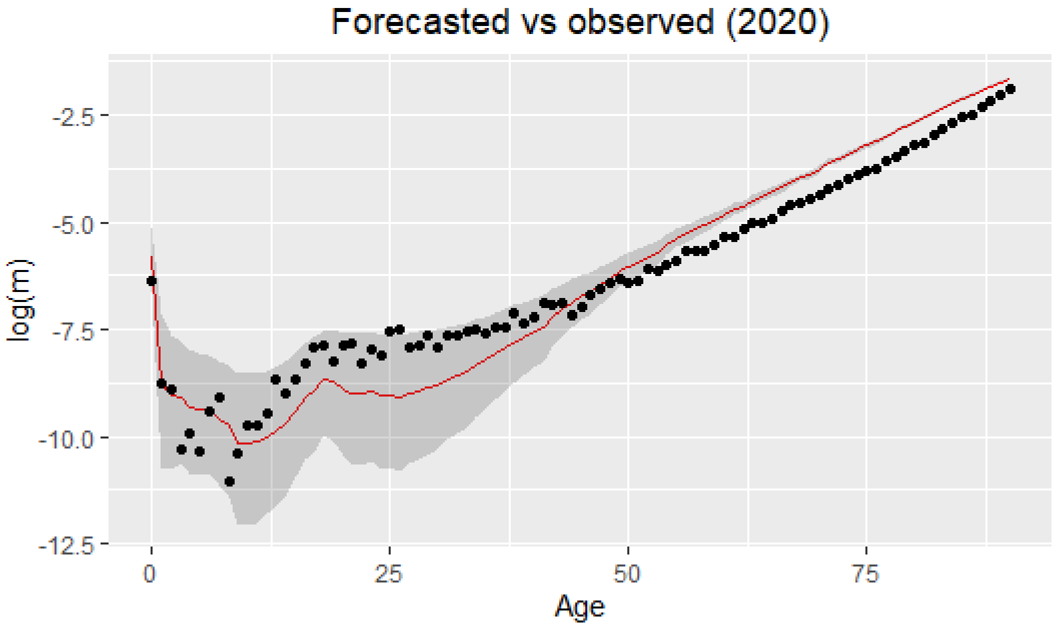

Figure 4.22: LC forecasted and empirical $\log(m)$ for 2020, with 95% CI.

implies equality between forecasted and historical.

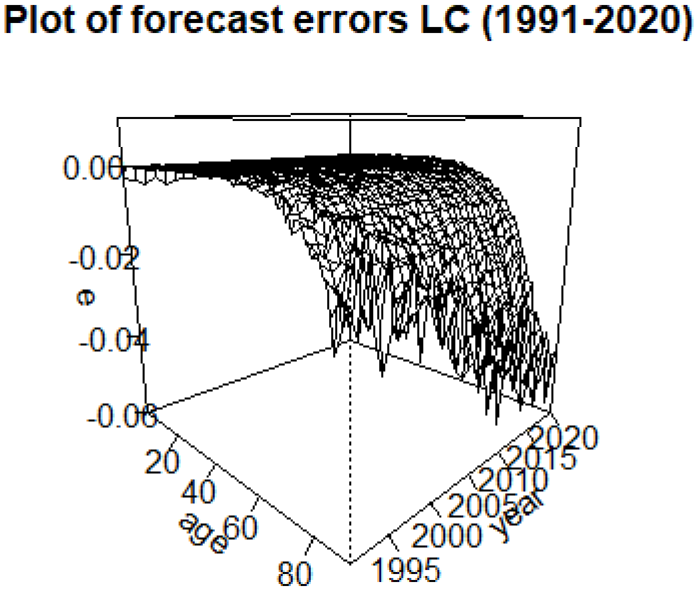

Figure 4.23: Forecast errors LC (1991 – 2020) (non log).

Figure 4.23 displays the forecast errors of the Lee-Carter predictions as a function of the age and year intervals. The errors are relatively small for younger ages and increases in magnitude towards the end of the age interval. From age 60 to 90 the errors are almost all negative, meaning that the Lee-Carter method overestimates the mortality rate for these ages. For the very high ages the errors also exhibit a significant increasing trend in magnitude over the year span: the longer the prediction interval, the greater the difference from actual data. We can conclude that the time trend decrease in mortality for high ages is not sufficiently captured in the model.

CHAPTER 5

Forecasting reserves

In life insurance, the insurer is concerned with the calculation of the future payments to the insured under the conditions of the contract given. The terms of a contract can include payments to the insured in case of illness/disability, unemployment, retirement, death etc. A main objective for the insurer is to compute the present values of the future cash flows and the mathematical reserve - the amount of money the insurance company has to have in reserve in order to stay solvent given the expected liabilities of the insurance policy. The uncertainty of such calculations arises from the state of the insured in the future as well as the interest rate behavior: at what probability will the insured be alive, unemployed, disabled, dead in the future? How will the interest rate evolve over time?

In this chapter we will only concern ourselves with pension and life insurance, where we are only interested in two states of the insured: alive or deceased. The statistical modeling of these will be based on the empirical mortality rates of the Norwegian population at year 2020 as well as the Lee-Carter projected mortality rates in section 4.5.

First, the concepts and definitions from probability theory required to calculate the present values will be detailed.

5.1 Definitions

All the definitions listed in this section are sourced from the book *"Stochastic Models in Life Insurance"* by Michael Koller.[Kol]

We let $(\Omega, \mathcal{A}, P)$ denote a *probability space* which satisfies Kolmogorov's axioms (listed below.)

1. $P(A) \in \mathbb{R}_+, \qquad \forall A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_1)$ for any countable sequence of disjoint sets $A_1, A_2, \ldots$

where $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ in axiom 1. Now, let $(S, \mathcal{S})$ be a measurable state space, that is S is a set and $\mathcal{S}$ is a σ-algebra on S, and denote by $T \subseteq \mathbb{R}$ a set. We then define:

Definition 5.1.1. *(Stochastic process)* A family of random variables $\{X_t : t \in T\}$ where

$$X_t : (\Omega, \mathcal{A}, P) \to (S, \mathcal{S}), \omega \mapsto X_t(\omega) \tag{5.1}$$

is called a *Stochastic process* on the probability space $(\Omega, \mathcal{A}, P)$ with state space S and parameter set T. The sample paths of the process are given by the function

$$X_{\cdot}(\omega) : T \to S, t \mapsto X_t(\omega).$$

Henceforward, we assume that each sample path is right continuous and have left limits P-a.s.

Definition 5.1.2. *(Indicator function with respect to a stochastic process)* Let $\{X_t\}_{t \in T}$ be a stochastic process as defined in definition 5.1.1. For $j \in S$, we define the indicator function with respect to the process $\{X_t\}_{t \in T}$ at time t as

$$\mathbb{I}_j(\omega) := \begin{cases} 1, & \text{if} X_t(\omega) = j, \\ 0, & \text{if} X_t(\omega) \neq j. \end{cases} \tag{5.2}$$

Markov Chain

In this section we go through definitions and results for a Markov Chain. In all preceding notations, the state space S is countable, that is $S = \mathbb{N}$.

Definition 5.1.3. *(Markov chain)* Let $\{X_t\}_{t \in T} \in S$ be a a stochastic process on $(\Omega, \mathcal{A}, P)$ with state space S and parameter set $T \in T \subseteq \mathbb{R}$ as defined in definition 5.1.1. Then X_t is a Markov chain if

$$P[X_{t_{n+1}} = i_{n+1} \mid X_{t_1} = i_1, X_{t_2}, \ldots, X_{t_n} = i_n] = P[X_{t_{n+1}} = i_{n+1} \mid X_{t_n} = i_n] \tag{5.3}$$

for all $t_1 < t_2 < \cdots < t_{n+1} \in T$, $i_1, i_2, \ldots, i_{n+1} \in S$, where $n \geq 1$, with $P[X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n] > 0$.

Remark: Equation 5.3 says that the process at time t_{n+1} only depends on the last state $X_{t_n} = i_n$; i.e. the probabilities does not depend on the path the process took to get to the last state i_n. For this reason, a Markov chain is often called a *memoryless* process.

Definition 5.1.4. *(Transition probability)* Let $\{X_t\}_{t \in T}$ be a stochastic process on $(\Omega, \mathcal{A}, P)$. The transition probabilities are defined as

$$p_{ij}(s, t) := P[X_t = j \mid X_s = i], \qquad s \leq t, \quad i, j \in S. \tag{5.4}$$

I.e. $p_{ij}(s, t)$ is the probability that the process X will switch i at time s to state j at time t.

Theorem 5.1.5. *(Chapman-Kolmogorov equation) Let $\{X_t\}_{t\in T}$ be a Markov chain as defined in definition 5.1.3 with transition probabilities $p_{ij}(s,t)$ as in equation 5.4. Then the following equation holds:*

$$p_{ij}(s,t) = \sum_{k\in S} p_{ik}(s,u)p_{kj}(u,t), \tag{5.5}$$

for all $s \leq u \leq t \in T$ and $i, j \in S$ with $P[X_s = i]$,$P[X_t = j] \neq 0$. Or equivalently, written on matrix form with notation $P(s,t)$

$$P(s,t) = P(s,u) \times P(u,t), \qquad s \leq u \leq t.$$

Theorem 5.1.6. *(Characterization of Markov chains) A stochastic process $\{X_t\}_{t\in T}$ is a Markov chain, if and only if*

$$P[X_{t_1} = i_1, \ldots, X_{t_n} = i_n] = P[X_{t_1} = \prod_{k=1}^{n-1} p_{i_k i_{k+1}}(t_k, t_{k+1}), \tag{5.6}$$

for all $t_1 < t_2 < \cdots t_n \in T$, $i_1, \ldots i_n \in S$, where $n \geq 1$.

5.2 The Insurance Model

As previously mentioned, our insurance policies will involve only two states of the insured: alive or dead. We let $S = \{*, \dagger\}$ denote the set of possible states, where $*$ is the state of being *alive* and $\dagger$ is the state of being *deceased*. Our survival model will be modeled by a *discrete time* Markov chain as defined in definition 5.1.3. We let $X = \{X_n, n \geq 0\}$ denote our Markov chain on a complete probability space $(\Omega, \mathcal{A}, P)$. From equation 5.4 we then have the two transition probabilities:

$$p_{**}(n,m) = P[X_m = * \mid X_n = *] \tag{5.7}$$
$$p_{*\dagger}(n,m) = P[X_m = \dagger \mid X_n = *], \tag{5.8}$$

where $n \leq m$ are discrete time points. The first equation is the probability of being alive at time m given that you are alive at time n, while the second equation denotes the probability of transitioning from the state of being alive in time n to the state of deceased at time m - i.e. the probability of dying during the discrete time interval $[n,m]$. Figure 5.1 displays our survival model at time n with the associated transition probabilities for the two states (obviously, if you are dead at time n the probability of staying dead is 1).

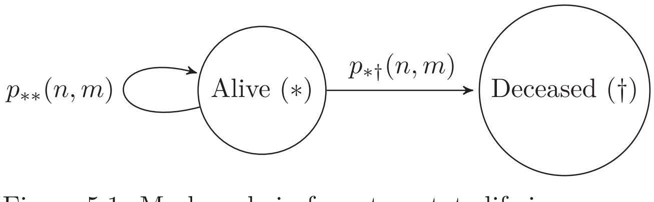

Figure 5.1: Markov chain for a two-state life insurance.

Figure 5.2 shows an example of mortality trajectory for a two-state model: An individual of age 25 sings an insurance contract and dies at the age of 55, 30 years later.

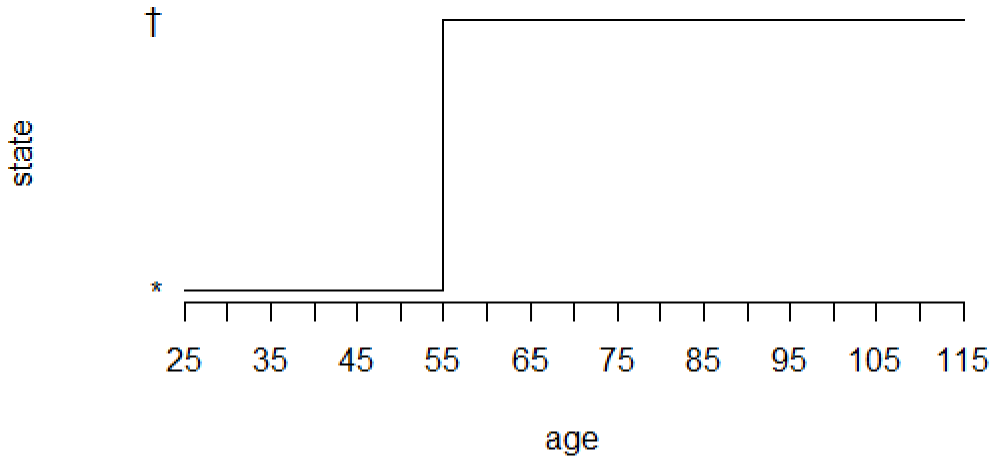

Figure 5.2: Trajectory of mortality in a two-state model

Transition probabilities To calculate the transition probabilities for future years, we use the Lee-Carter forecasted mortality rates $\tilde{m}_{x,t} = \tilde{\mu}_{x,t}$ for ages $0-90$ and years $2021-2060$ obtained in chapter 4, section 4.5. For an individual age x in year t (the beginning of the contract) we denote the mortality at year $t+s$ as follows:

$$\mu^x_{*\dagger}(s) := \tilde{\mu}_{x+s,t+s}. \tag{5.9}$$

The yearly transition probabilities are then derived from equation 2. We have:

$$p^x_{*\dagger}(s,s+1) := p_{*\dagger}(x+s,x+s+1) = 1-\exp(-\mu^x_{*\dagger}(s)) = 1-\exp(-\tilde{\mu}_{x+s,t+s}) \tag{5.10}$$

$$p^x_{**}(s,s+1) := p_{**}(x+s,x+s+1) = \exp(-\tilde{\mu}_{x+s,t+s}), \tag{5.11}$$

where x is the age of the individual at the beginning of the insurance contract and s is the number of years into the contract; thereby making $x+s$ the age of the insured at year $t+s$. The result of the survival probability $p^x_{**}(s,s+1)$ follows from the complementary of the probability of mortality.

Furthermore, from the definition of the Chapman-Kolmogorov equation (5.5), we expand our survival probability to $s+2$ years into the contract and get

$$\begin{aligned}
p^x_{**}(s,s+2) &= \sum_{k\in S} p^x_{*k}(s,s+1)p^x_{k*}(s+1,s+2)\\
&= p^x_{**}(s,s+1)p^x_{**}(s+1,s+2) + p^x_{*\dagger}(s,s+1)p^x_{\dagger *}(s+1,s+2)\\
&= p^x_{**}(s,s+1)p^x_{**}(s+1,s+2) = \exp(-\tilde{\mu}_{x+s,t+s})\exp(-\tilde{\mu}_{x+s+1,t+s+1})\\
&= \exp(-\tilde{\mu}_{x+s,t+s} - \tilde{\mu}_{x+s+1,t+s+1})
\end{aligned}$$

where the last equation is a result of $p^x_{\dagger *}(s+1,s+2)=0$, since death is an absorbing state of no return. By the same procedure, and by using the complementary condition of mortality and survival in a two-state model, we arrive at the probability of transitioning to death within $s+2$ years into the contract

$$\begin{aligned}
p^x_{*\dagger}(s,s+2) &= p^x_{s,s+1}(*\dagger) + p^x_{**}(s,s+1)p^x_{*\dagger}(s+1,s+2)\\
&= 1-\exp(-\tilde{\mu}_{x+s,t+s}) + \exp(-\tilde{\mu}_{x+s,t+s})(1-\exp(-\tilde{\mu}_{x+s+1,t+s+2})),
\end{aligned}$$

with the interpretation that the probability of dying within two discrete time periods (years) can only happen at either two points in time: either the individual dies at time $s+1$ or the individual survives to time $s+2$ and then dies, ergo the sum of these two cases make the probability.

Since the Chapman-Kolmogorov equation is valid for all integers $n>s$, we expand and get the general expressions for the transition probabilities at a future time

$$p^x_{**}(s,n) = \prod_{i=s}^{n} p^x_{**}(i,i+1)$$

$$p^x_{*\dagger}(s,n) = p^x_{*\dagger}(s,s+1) + \sum_{k=s+2}^{n} p^x_{*\dagger}(k-1,k)\prod_{j=s}^{k-2} p^x_{**}(j,j+1).$$

(Note that these results are valid for any $s\geq 0$, so these equations can be used from the year the contract is entered $s=0$)

5.3 Policy functions

We define two types of insurance policy functions in discrete time, a^{pre}_i and a^{post}_{ij}:

- $a^{pre}_i(n) =$ payment to the insured at time n, given that the insured is in state i at time n
- $a^{post}_{ij}(n) =$ capital benefits for switching from state i to state j at time $n+1$,

where *pre* means that the payment is transferred at time n and *post* means that the payment is transferred at the end of the interval $[n, n+1)$, at time $n+1$. We will use negative sign for payments going from the policy holder to the insurance company, and positive sign for the converse case. For states $i, j \in S$ the stochastic cash flow of an insurance is then given by the equation

$$A(s) := \sum_{i \in S} A_i^{pre}(s) + \sum_{\substack{i,j \in S \\ j \neq i}} A_{ij}^{post}(s), \tag{5.12}$$

for every integer $s \geq 0$, and

$$A_i^{pre}(s) := \sum_{n \geq 0}^{s} \mathbb{I}_{\{X_n = i\}} a_i^{pre}(n), \qquad A_{ij}^{post}(s) := \sum_{n=0}^{s} \mathbb{I}_{\{X_n = i, X_{n+1} = j\}} a_{ij}^{post}(n). \tag{5.13}$$

for $n = 0, 1 \ldots$. The variations of A happen at discrete time intervals $[s, s+1)$, $s = 0, 1, \ldots$ and are written as[Kol]

$$\Delta A_i^{pre}(s) = \mathbb{I}_i(s) a_i^{pre}(s), \tag{5.14}$$

$$\Delta A_{ij}^{post}(s) = \Delta N_{ij}(s) a_{ij}^{post}(s), \tag{5.15}$$

$$\Delta A(s) = \sum_{i \in S} \Delta A_i^{pre}(s) + \sum_{i,j \in S} \Delta A_{ij}^{post}(s), \tag{5.16}$$

where $\Delta N_{ij}(s)$ denotes the number of jumps from state i to j in the time interval $(0, s)$.

5.4 Calculating reserves

In this section we go through the necessary steps to calculate the forecasted reserves for life- and pension insurance based on Norwegian mortality rates. To calculate the reserves we need to first calculate the present value of the cash flows and liabilities A. From the book of Koller we have (with our notation) that the *prospective value* of a stochastic cash flow A is defined as

Definition 5.4.1. Let $V^+(s, A)$ denote the prospective value of a stochastic cash flow A at discrete time s. It is then defined as

$$V^+(s, A) := \frac{1}{v(t)} \left[\sum_{i \in S} \sum_{n=s}^{\infty} v(n) \Delta A_i^{pre}(n) + \sum_{i,j \in S} \sum_{n=s}^{\infty} v(n+1) \Delta A_{ij}^{post}(n) \right], \quad s \in \mathbb{N}, \tag{5.17}$$

where $v(s)$ is called *discount factor* and is defined as

$$v(s) := \exp\left(- \int_0^s r_u du \right) \tag{5.18}$$

where $r : [0, \infty) \to \mathbb{R}$ is a deterministic and integrable function which models the interest rate.

We now define the explicit formula for the prospective value of the mathematical *reserves* which is given by the conditional expectations of the prospective value of stochastic cash flow[Kol]

Theorem 5.4.2. *Let x be the age of the age of the insured at the beginning of the contract. The value of the liability A at discrete time s given that the insured is in state i at time s i given by*

$$V_i^+(s,A) := E[V^+(s,A) \mid X_s = i] \tag{5.19}$$

$$= \frac{1}{v(s)} \left[\sum_{j\in S} \sum_{n\geq s} v(n) p_{ij}^x(s,n) a_j^{pre}(n) + \sum_{\substack{j,k\in S\\ k\neq j}} \sum_{n\geq s} v(n+1) p_{ij}^x(s,n) p_{jk}^x(n,n+1) a_{jk}^{post}(n) \right]. \tag{5.20}$$

The explicit solution in equation 5.20 is numerically intensive to calculate. The solution can be obtained simpler by Thiele's difference equation, which is a recursive equation for discrete time n, and with our notation defined as[Kol]

Theorem 5.4.3. *(Thiele's difference equation).*

$$V_i^+(n) = a_i^{pre}(n) + \sum_{j\in S} v_n p_{ij}^x(n,n+1)\left(a_{ij}^{post}(n) + V_j^+(n+1)\right), \quad n \in \mathbb{N} \tag{5.21}$$

where $v_t = \frac{1}{1+r_n}$ is our notation for the one year/time discount factor. The recursion works by finding $V_i^+(n-1), \ldots V_i^+(0)$ when the final amount $V_i^+(T)$ is known (where T is the end of the contract.)

In the next sections, when calculating the reserves for different insurance scenarios, we will be using Thiele's difference equation programmed in R.

5.5 Life insurance (endowment)

Let us consider an example of an Norwegian individual age 30 in year 2021. The individual signs a life insurance contract with the following specifications: The insurance contract ends in 37 years when the individual is of Norwegian pension age, that is 67 years old, in year 2058. We denote by $T = 37$ the length of the contract and set the yearly interest rate $r = 3\%$. The insured is guaranteed NOK 1 000 000 (Norwegian crowns) if he/she survives to age 67 or else the insured receives NOK 2 000 000 in case of death during the contract period. We then have the discrete policy functions

$$a_*^{pre}(n) = \begin{cases} 0, \ n = 0, \ldots, 36, \\ 1\,000\,000 \ n = 37 \end{cases}, \quad a_{*\dagger}^{post}(n) = \begin{cases} 2\,000\,000, \ n = 0, \ldots, 36, \\ 0 \text{ otherwise} \end{cases} \tag{5.22}$$

$$\tilde{a}_*^{pre}(n) = \begin{cases} -\pi, \ n = 0, \ldots, 36, \\ 0, \ n = 37 \end{cases} \tag{5.23}$$

where π is the yearly premiums the insured has to pay for the policy. The premium will be calculated using the *equivalence principle*, such that the cost of the premium is zero at the start of the contract ($V_*^+(0, A) = 0$). We will also look at the case where this is paid as a single lump sum at year 0, then denoted π_0.

Using Thiele's difference equation (5.21) for these policy functions we get present value for endowment equals

$$V_*^+(n, A_*^{pre}) = a_*^{pre}(n) + v_n p_{**}^x(n, n+1) V_*^+(n+1, A_*^{pre})$$

the present value for the death benefits equals

$$V_*^+(n, A_{*\dagger}^{post}) = v_n(p_{**}^x(n, n+1) V_\star^+(n+1, A_{*\dagger}^{post}) + p_{*\dagger}^x(n, n+1) a_{*\dagger}^{post}(n)$$

and the present value for the premiums

$$V_*^+(n, \tilde{A}_*^{pre}) = -\pi + v_n p_{**}^x(n, n+1) V_*^+(n+1, \tilde{A}_*^{pre})$$

where in all equations $v_n = \frac{1}{1+r_n}$ is the one year time-discount factor

.

Using R, we calculate the transition probabilities using the Lee-Carter forecasted $\tilde{m}_{x,t}$ for all the relevant years $2021 - 2058$. The function for doing so is included in the snipped below, together with the function for the policy functions.

```
p_aa <- function(age, year){
  p_surv <- exp(-exp(M.tilde[age+1, year]))
  return(p_surv)
}

# policy functions
a_pol <- function(age){
  a <- rep(0, 2)
  if (age == 66){a[1] = 1000000}
  if (age < 67){a[2] = 2000000}
  return(a)
}
```

Then we find $V_i^+(n-1), \ldots, V_i^+(0)$ with the code below:

```
r = 0.03
res_endow <- rep(0, 38)
res_death <- rep(0, 38)
V_prem <- rep(0, 38)
year_set <- 37:1
# Thiele's difference equation
for (i in 1:37){
  res_endow[38-i] <- exp(-r)*p_aa(67-i, year_set[i])*(a_pol(67-i)[1]
                                                      +res_endow[38-i+1])
  res_death[38-i] <- exp(-r)*(p_aa(67-i, year_set[i])
                              *res_death[38-i+1]+(1-p_aa(67-i, year_set[i]))
                              *(a_pol(67-i)[2]))
  V_prem[38-i] <- -1 + exp(-r)*p_aa(67-i, year_set[i])*V_prem[38-i+1]
}
res_endow[38] <- 1000000
res_total <- res_endow + res_death
pi_0 <- res_total[1]
pi <- -(pi_0/V_prem[1])
```

The result of the procedure is shown in table 5.1. Observe that the present value for endowment is increasing for every year of age, which makes sense since for every increasing year the probability that the insured will be alive for the payout increases. The present value for death benefits increases until age 48 and thereby decreases until the end of the contract time. From the present value of the total at age 30, i.e. the year the contract begins, we see that the contract will cost the insurance company NOK 367239.08. This is the amount of money the insurance company will charge if it chooses to charge a lump sum premium π_0.

Calculating yearly premiums

Under the *equivalence principle*, we want chose π such that $V_*^+(0, A) = 0$. I.e. we want to chose yearly premiums such that cost of the insurance is 0 at the beginning of the contract

$$0 = V_*^+(0, \tilde{A}_*^{pre}) + V_*^+(0, A_*^{pre}) + V_*^+(0, A_{*\dagger}^{post}).$$

But the question remains as to what the value of the yearly premium π is. By setting $\pi = 1$ in an artificial policy $V_*^+(n, \tilde{A}_0^{prem=1})$, we get that Thiele's equation is simply

$$V_*^+(n, \tilde{A}_*^{prem=1}) = -1 + v_n p_{**}(n, n+1) V_*^+(n+1, \tilde{A}_*^{prem=1}), \quad n = 0, 1, \ldots, 36$$

with end of contract condition $V_*^+(37, \tilde{A}_*^{prem=1}) = 0$ as described in the policy. We then want to find π such that

$$\pi V_*^+(0, \tilde{A}_*^{prem=1}) + V_*^+(0, A_*^{pre}) + V_*^+(0, A_{*\dagger}^{post}) = 0. \tag{5.24}$$

using R, we obtain

$$V_*^+(0, \tilde{A}_*^{prem=1}) = -22.39019, \tag{5.25}$$

and hence

Age	PV endowment	PV death benefit	PV Total
30	309300.96	57938.12	367239.08
31	318865.30	58821.54	377686.84
32	328748.66	59594.97	388343.63
33	338937.17	60399.19	399336.35
34	349437.13	61252.20	410689.33
35	360268.22	62100.17	422368.39
36	371439.26	62952.38	434391.64
37	382957.49	63826.95	446784.44
38	394850.54	64642.18	459492.72
39	407102.43	65532.60	472635.04
40	419773.51	66270.95	486044.46
41	432833.47	67056.84	499890.31
42	446357.25	67618.27	513975.52
43	460270.74	68335.05	528605.79
44	474649.06	68947.53	543596.59
45	489493.92	69510.59	559004.51
46	504797.36	70113.06	574910.42
47	520680.86	70357.79	591038.64
48	537107.30	70455.11	607562.41
49	554140.66	70246.64	624387.31
50	571720.29	70011.03	641731.31
51	589958.32	69438.39	659396.71
52	608848.25	68625.11	677473.35
53	628373.38	67692.41	696065.79
54	648779.32	65970.66	714749.98
55	669799.98	64331.58	734131.55
56	691901.07	61520.50	753421.57
57	714457.33	59361.18	773818.51
58	738069.25	56287.92	794357.17
59	762421.50	53215.25	815636.74
60	788027.48	48926.45	836953.93
61	814412.50	44687.93	859100.44
62	841759.32	40125.06	881884.38
63	870823.11	33605.40	904428.52
64	901057.14	26495.47	927552.61
65	932380.24	19055.01	951435.25
66	965596.86	9697.34	975294.20
67	**1000000**	**0**	**1000000**

Table 5.1: Reserves for life insurance (endowment), $r = 3\%$, $x = 30$, $T = 37$

$$\pi = -\frac{367239.08}{-22.39019} = 16401.79$$

The present value of the premium, the present value of the benefits (insurance cost) and the mathematical reserves are plotted in figure 5.3.

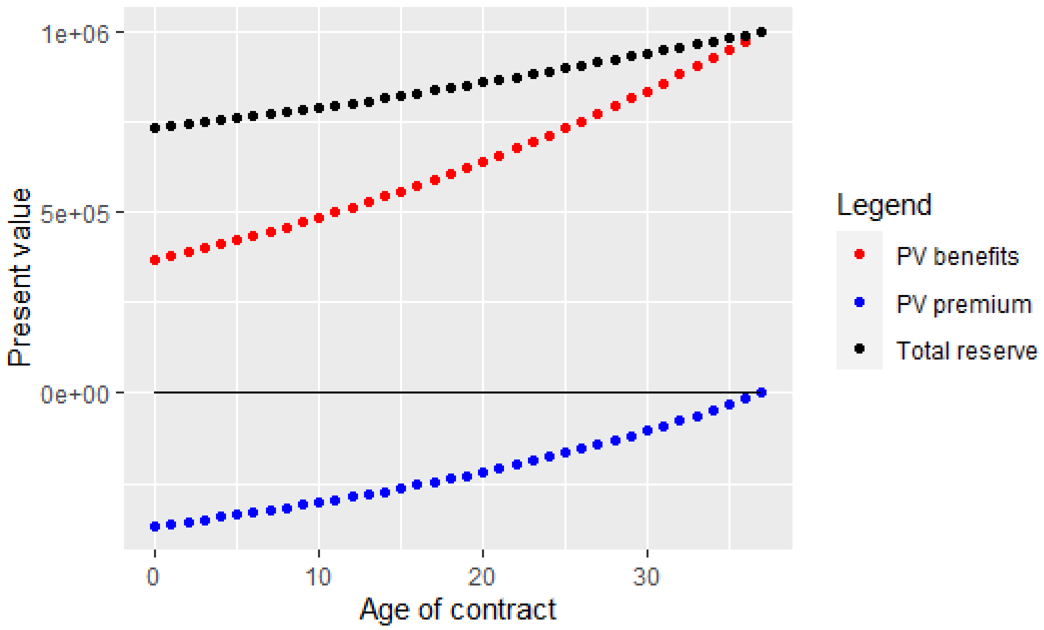

Figure 5.3: Present value endowment policy $r = 3\%$, $x = 30$, $T = 37$.

PV different interest rates

Figure 5.6 shows the present value for total reserves for different interest rates r. For lower interest rates, the present value is higher, which means that the cost of the contract is higher for the insurance company. If the company decides to charge a lump sum premium π_0 at the beginning of the contract, this amount is NOK 887822.2 for the low interest rate $r = 0.5\%$ and NOK 5766.877 for the high interest rate of 20%.

5.6 Pension insurance

Let us now consider a pension insurance. A Norwegian individual age 30 signs a contract in year 2000 which grants him/her a yearly pension of NOK 130 000 from the retirement age of 67 up to (but not including) age 90. The insured pays a yearly premium π up until the age of retirement. n denotes the age of the contract. In the first part of the analysis, we set the interest rate $r = 3\%$. The mortality rates from the Lee-Carter estimation in section 4.4 will be used for years $2000 - 2020$, while the forecasted mortality rates from section 4.5 will be used for the years we do not have historic data, namely $2021 - 2060$. The

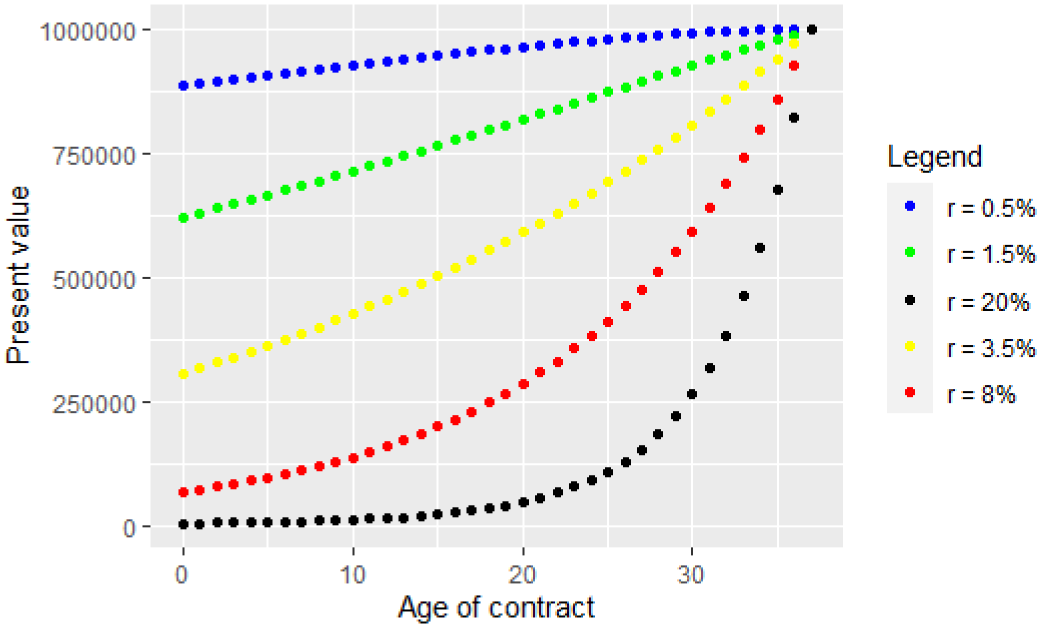

Figure 5.4: Present value total reserves for endowment where $r = 0.5\%, 1.5\%, 3.5\%, 8\%, 20\%$.

policy functions for this contract are given by

$$a_*^{pre}(n) = \begin{cases} 130\,000, & n = 37, \dots, 59, \\ 0 \text{ otherwise} \end{cases} \quad , \quad \tilde{a}_*^{pre}(n) = \begin{cases} -\pi, & n = 0, \dots, 36, \\ 0 \text{ otherwise} \end{cases} \tag{5.26}$$

Which give the following Thiele's equations for present value of benefits and present value of premiums

$$V_*^+(n, A_*^{pre}) = a_*^{pre}(n) + v_n p_{**}^x(n) V_*^+(n+1, A_*^{pre}),$$
$$V_*^+(n, \tilde{A}_*^{pre}) = \tilde{a}_*^{pre}(n) + v_n p_{**}^x(n) V_*^+(n+1, \tilde{A}_*^{pre})$$

As for the life insurance in the previous section, we solve Thiele's difference equations in R. As before, the equation for the premium is calculated with an artificial policy where $\pi = 1$, but this time, since the insured only pays premiums up until the age of retirement, we have

$$\tilde{a}_*^{prem=1} = \begin{cases} -1, & n = 0, \dots, 36, \\ 0 \text{ otherwise.} \end{cases} \tag{5.27}$$

The calculations yield $\pi_0 = 540819.9$ and yearly premiums $\pi = 24273.45$. The present value for the benefits, the present value for premium and the mathematical reserves for select age x (age of insured) are tabulated in table 5.2. The full set of values are plotted in figure 5.5. The present values of the premiums are slowly decreasing in absolute value until they zero out at age 67. The present value for the benefit - the total cost of the insurance - reaches its peak at retirement age 67, from whence it is equal to the mathematical reserve and slowly declines towards zero at age 90, when the insured no longer receives a pension from the contract.

Age	PV benefits	PV premium	Mathematical reserve
30	540819.87	-540819.87	1081639.73
35	630495.30	-497293.86	1127789.16
40	735446.65	-446822.51	1182269.17
45	858807.56	-388434.05	1247241.61
50	1005250.93	-321108.93	1326359.87
55	1181486.28	-243502.39	1424988.67
60	1396163.21	-153365.06	1549528.27
65	1664445.24	-47693.96	1712139.19
70	1603394.85	0.00	1603394.85
75	1264343.27	0.00	1264343.27
80	898152.13	0.00	898152.13
85	492955.17	0.00	492955.17
86	407313.20	0.00	407313.20
87	315121.01	0.00	315121.01
88	219648.79	0.00	219648.79
89	**114859.58**	**0.00**	**114859.58**
90	0.00	0.00	0.00

Table 5.2: Reserves for pension insurance, $r = 3\%$, $x = 30$, $T = 60$

PV different interest rates

Figure 5.3 shows the present value of the pension benefits for select interest rates 1%, 2%, 3%, 8% and 20%, while figure 5.7 shows the mathematical reserves for the same interest rates. The higher the interest rate, the lower cost of the policy from the insurance company's standpoint, and the less money needed in reserves. The mathematical reserves needed at the start of the contract when the interest rate is 1% is NOK 2770048, while the amount needed for an interest rate of 20% is only NOK 613.1975. Even an increase from $r = 1\%$ to $r = 2\%$ decreases the amount with NOK 959271 to NOK 1727544, showing the impact interest rates have for long-term contracts like this.

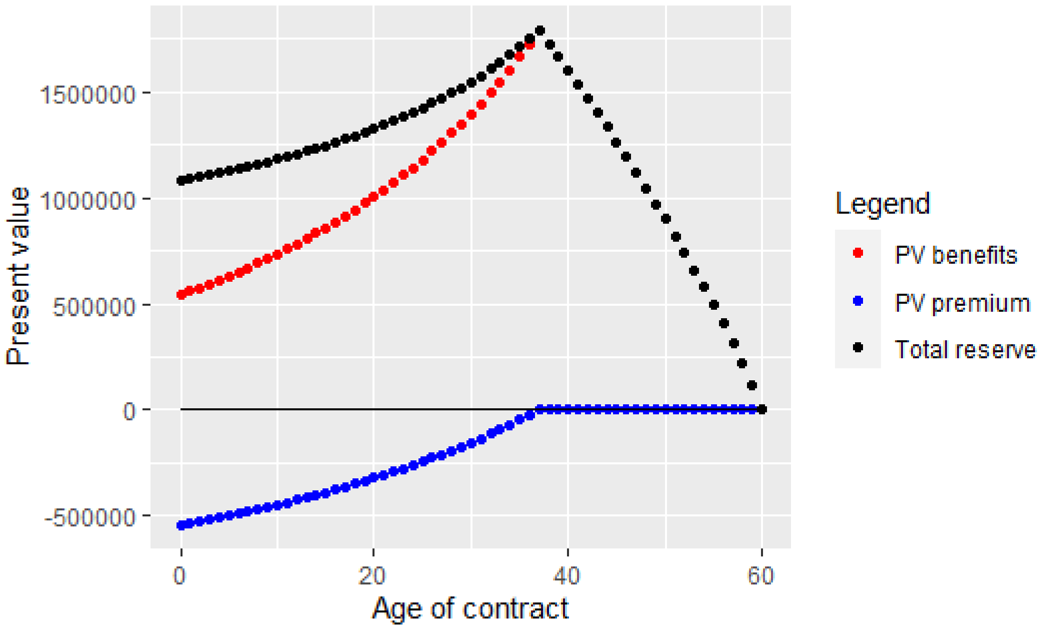

Figure 5.5: Present value pension insurance, $r = 3.5$, $x = 30$, $T = 60$.

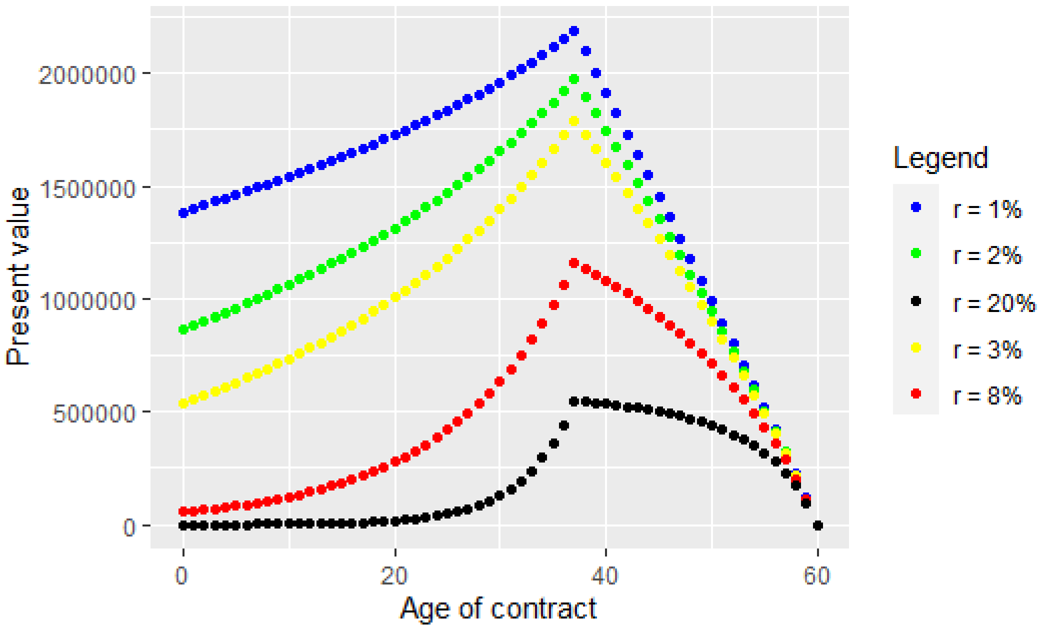

Figure 5.6: Present value of pension benefits, $r = 1\%, 2\%, 3\%, 8\%, 20\%$.

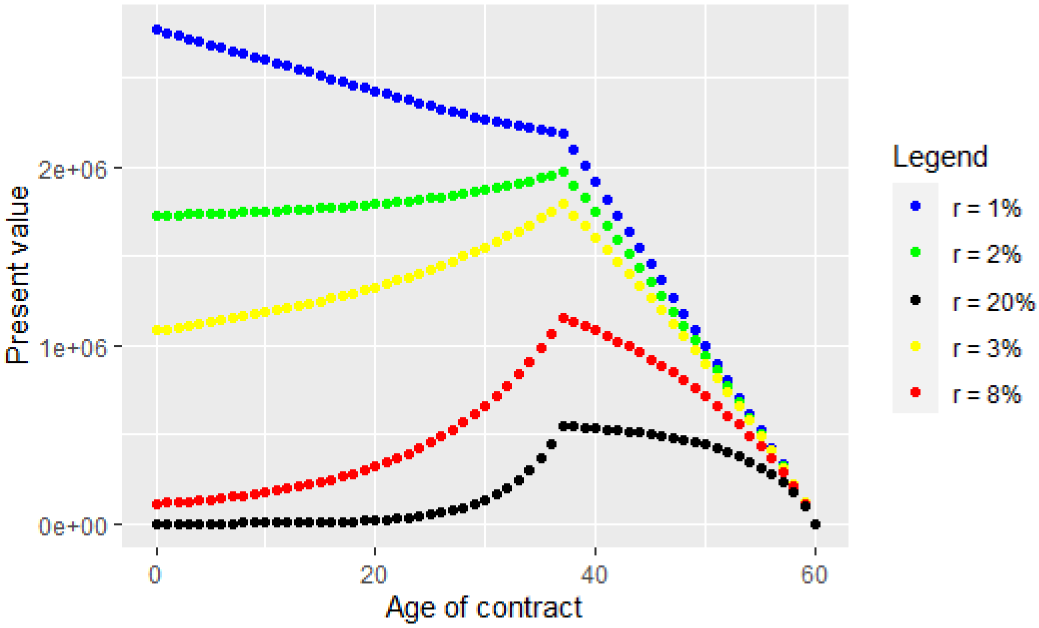

Figure 5.7: Mathematical reserves pension insurance, $r = 1\%, 2\%, 3\%, 8\%, 20\%$.

CHAPTER 6

Conclusion

The conclusion and summary of the thesis is presented in this chapter.

Conclusive remarks

In this thesis the standard Lee-Carter model has been fitted on historical both-gendered mortality data for the Norwegian population. The estimation has then been assessed based on the models ability to capture both age and time trends of the empirical data, as well as the ratio of variance it is able to explain for different ages grouped together. From these investigations we saw that the overall trends were captured sufficiently. However, the model was only able to explain $\approx 50\%$ of the variance for ages $21 - 30$. This is an age group that has had a lot of variation over the years of analysis $1960 - 2020$, thus not necessarily a flaw in the model. To more accurately fit this subset of the population, a more advanced model might be needed, and more historical data might benefit the estimation. For higher ages, which are more relevant from a pension insurers perspective, the estimations followed the empirical data with great accuracy.

The Lee-Carter forecasts for years $2021 - 2060$ followed the general historical time trend of decreasing mortality rates for all ages $0 - 90$. One question that arises is if this decreasing trend is realistic in the next 40 years or not. It is not clear if this will continue, especially not for all ages. This will always be an inherent flaw when making predictions made on previous values. Comparing the Lee-Carter forecasts with the ones proposed in K2013, both models yielded very similar results for low to middle ages. For high ages, the Lee-Carter model predicted increasingly higher mortality rates than K2013.

To measure the credibility of forecasts made by the Lee-Carter model on Norwegian mortality, two different year partitions were made to create one training set and one test set. The partition chosen for the training set included the years of World War 2, which is what you might call a highly unusual event which caused a spike in Norwegian mortality. The resulting prediction errors might have suffered as an effect. The Lee-Carter model overestimated the mortality rates for higher ages, compared to the test set. This was reflected in the forecast errors, which where increasing in magnitude the higher the ages predicted; inaccuracies which increased over future forecasted years. Long-term forecasts for pension ages are therefore non-satisfactory based on the years this model was fitted on.

The Lee-Carter forecasted mortality rates for years $2021-2060$, predicted based empirical data for years $1960-2020$, were used to compute present values and reserves for an endowment policy and a pension policy. It was also demonstrated how to compute the yearly premiums, present values and mathematical reserves of the contracts. If the predicted mortality rates these computations are based on are accurate remains to be seen, though they were found to be fairly similar to the ones proposed by the financial supervisory authority of Norway in K2013. Though, the fact that the Lee-Carter forecasted rates were higher - and even more so for higher ages - than the ones produced from K2013, has an impact on the calculations. The choice between these methods will affect the calculation of reserves and premiums in pension and life insurance, especially for long-term contracts which rely on predictions far into the future. It is of importance for insurers and policy makers to accurately compute costs associated with policies.

Future work

It remains to see if K2013 and Lee-Carter forecasts differ substantially for each gender separate. Another interesting question is the real effect these differences have on insurance policies. With regards to Lee-Carter fitted on the year interval $1920-1991$, more analysis need to done. Various extensions of the Lee-Carter model might have more predictive power when fitted on time intervals with such highly unusual events as a world war.

Regarding covid, the analysis performed in this thesis could not find a substantial increase in mortality in year 2020, even for ages more exposed to severe complications from the virus. In this regard, interesting studies into cause-specific mortality rates could be performed, not only for 2020 but also 2021 and 2022, to better understand the effect of the pandemic.

Appendices

APPENDIX A

The First Appendix: Proofs

Relationship in equation 2.6

Proof. Using the relationship between a distribution function and density function of a random variable from probability theory

$$F(t) = \int_0^t f(s)ds,$$

we want to find the density function $f_x(t)$ of the time variable T_x. From the definition of derivation we have

$$\begin{aligned} f_x(t) = \frac{\partial}{\partial t} F_x(t) &= \lim_{\Delta t \to 0^+} \frac{F_x(t+\Delta t) - F_x(t)}{\Delta t} \\ &= \lim_{\Delta t \to 0^+} \frac{P[T_x \leq t+\Delta t] - P[T_x \leq t]}{\Delta t} \\ &= \lim_{\Delta t \to 0^+} \frac{P[T_x \leq t] + P[T_x > t] P[T_x \leq t+\Delta t \mid T_x > t] - P[T_x \leq t]}{\Delta t} \\ &= P[T_x > t] \lim_{\Delta t \to 0^+} \frac{P[T_x \leq t+\Delta t \mid T_x > t]}{\Delta t}, \end{aligned}$$

where in the second step we have used the definition of the lifetime distribution in equation 2.1, and in the third step we have used that the time series T_x is increasing. Now, we observe that the expression we ended up with is just the survival function $S_x(t)$ and the hazard function μ_{x+t}, hence

$$\frac{\partial}{\partial t} F_x(t) = S_x(t)\mu_{x+t}.$$

Since $F_x(t) = 1 - S_x(t)$, we can write this differential equation as

$$\frac{\partial}{\partial t} S_x(t) = -S_x(t)\mu_{x+t},$$

which is an ordinary differential equation we can solve with boundary condition $S_x(0) = 1$ as follows:

$$\begin{aligned}\frac{\partial}{\partial t}S_x(t) = -S_x(t)\mu_{x+t}\\ \Rightarrow \frac{\partial}{\partial t}\log(S_x(t)) = -\mu_{x+t}\\ \Rightarrow \int_0^t \left(\frac{\partial}{\partial s}\log(S_x(t))\right) ds = -\int_0^t \mu_{x+s} ds + c\\ \Rightarrow \log(S_x(t)) = -\int_0^t \mu_{x+s} ds + c.\end{aligned}$$

And since the boundary condition implies that the integration constant $c = 0$, the results follows by taking the exponential:

$$S_x(t) = \exp\left(-\int_0^t \mu_{x+s} ds\right).$$

APPENDIX B

The Second Appendix: R-code

```r
# 1) Data handling

# Extract data
deaths <- read.table("NOR_deaths.txt", header = TRUE, skip = 2)
expos <- read.table("NOR_exposures.txt", header = TRUE, skip = 2)
deaths$Year <- as.numeric(deaths$Year)
expos$Year <- as.numeric(expos$Year)

# Split data into female, male, both
Male.deaths <- deaths$Male
Male.expos <- expos$Male
Female.deaths <- deaths$Female
Female.expos <- expos$Female
Both.deaths <- deaths$Total
Both.expos <- expos$Total

# Matrix forms of deaths and exposure-to-risk
Age <- 0:110
Year <- 1846:2020
Deaths <- matrix(Both.deaths, nrow = 111)
Expos <- matrix(Both.expos, nrow = 111)
colnames(Deaths) <- Year
rownames(Deaths) <- Age
colnames(Expos) <- Year
rownames(Expos) <- Age


# Get data from year 1960-2020 and ages 0-90
D <- Deaths[Age <= 90, (Year >= 1960) & (Year <= 2020)]
E <- Expos[Age <= 90, (Year >= 1960) & (Year <= 2020)]


# Surface plots of death counts and exposure-to-risk
persp(seq(0,90), seq(1960, 2020), D, theta = 40, phi = 25,
      ticktype = "detailed", xlab = "Age", ylab = "Year",
      zlab = "Deaths", col = "lightgrey",
      cex.lab = 0.8, cex.axis = 0.8)

persp(seq(0,90), seq(1960, 2020), E, theta = 45, phi = 15,
      ticktype = "detailed", xlab = "Age", ylab = "Year",
      zlab = "Exposures", col = "lightgrey",
      cex.lab = 0.8, cex.axis = 0.8)


# Function returning crude mort
```

```
get.Crude.M <- function(d, e){
  crude.m <- d/e
  for (i in 1:nrow(crude.m)){ # Replace zeros with mean last and next year
    for (j in 1:ncol(crude.m)){
      if (crude.m[i,j] == 0){
        crude.m[i,j] <- mean(c(crude.m[i,j-1],crude.m[i,j+1]))
      }
    }
  }
  return(crude.m)
}

# Crude mortality rate matrix
crude.mort <- get.Crude.M(D, E)

# Chapter 3: Gompertz for year 2000 (un-weighted least squares)
#Age 40-90
age <- 40:90
fit.gomp1 <- lm(log(crude.mort[41:91,41]) ~ age)
# data frame
gomp.df1 <- data.frame(age+0.5, log(crude.mort[41:91, 41]), fit.gomp1$fitted.values)
colnames(gomp.df1) <- c("age", "obs", "fit")

# plot
colors <- c("Observed" = "light blue", "Fitted" = "Red")
ggplot(data=gomp.df1, aes(x=age))+
  geom_point(aes(x = age, y = obs, color="Observed"))+
  geom_line(aes(x = age, y = fit, color="Fitted"))+
  labs(x = "Age", y = expression(paste("log(", mu, ")")), color="Legend",
       title = "Gompertz age 40-90, year 2000")+
  theme(plot.title = element_text(hjust = 0.5))

# Gompertz Age 0-90, year 2000
age <- 0:90
fit.gomp2 <- lm(log(crude.mort[,41]) ~ age)
summary(fit.gomp2)
# data frame
gomp.df2 <- data.frame(age+0.5, log(crude.mort[,41]), fit.gomp2$fitted.values)
colnames(gomp.df2) <- c("age", "obs", "fit")

colors <- c("Observed" = "light blue", "Fitted" = "Red")
ggplot(data=gomp.df2, aes(x=age))+
  geom_point(aes(x = age, y = obs, color="Observed"))+
  geom_line(aes(x = age, y = fit, color="Fitted"))+
  labs(x = "Age", y = expression(paste("log(", mu, ")")), color="Legend",
       title = "Gompertz age 0-90, year 2000")+
  theme(plot.title = element_text(hjust = 0.5))

################################################################

# Surface plot observed mortality
persp(seq(0,90)+0.5, seq(1960, 2020), log(crude.mort), theta = -45, phi = 30,
      ticktype = "detailed", xlab = "Age", ylab = "Year",
      zlab = "log(m)", col = "lightgrey",
      cex.lab = 0.8, cex.axis = 0.8)


# Line-plots observed mortality for high ages
color <- rainbow(3)
ageset <- c(80, 85, 90)
plot(1960:2020, crude.mort[ageset[1]+1,], type="l", col=color[1],
     ylim=c(0.0, 0.3), ylab = "Mortality rate", xlab = "Year")
```

```
legend("topright", legend = c("age 80", "age 85", "age 90"), cex = 0.71,
       lty = c(1,1,1), col = color)
for (i in 2:3){
  lines(1960:2020, crude.mort[ageset[i]+1,], col=color[i])
}

# Chapter 4: Covid-19 calculations
# High ages 80-90
cov.age <- 80:90
cov.mort20 <- c() # vector for mortality year 2020
cov.dif.1 <- c() # 1 year difference in percentage
cov.dif.2 <- c() # 2 year difference in percentage
cov.dif.3 <- c() # 3 year difference in percentage
for (i in 1:11) {
  cov.mort20[i] <- crude.mort[80+i, 61]
  cov.dif.1[i] <- ((cov.mort20[i] - crude.mort[80+i, 60])/cov.mort20[i])*100
  cov.dif.2[i] <- ((cov.mort20[i] - crude.mort[80+i, 59])/cov.mort20[i])*100
  cov.dif.3[i] <- ((cov.mort20[i] - crude.mort[80+i, 58])/cov.mort20[i])*100
}
cov.matrix <- cbind(cov.mort20, cov.dif.1, cov.dif.2, cov.dif.3)
colnames(cov.matrix) <- c(’age’, ’mortality’, ’1 yr diff %’, ’2 yr diff’,
                          ’3 yr diff’)
row.names(cov.matrix) <- cov.age
tab.matrix.cov <- as.table(cov.matrix)
tab.matrix.cov

# Lee-Carter algorithm function
get.LC <- function(M){
  ax <- rowMeans(log(M))
  A <- log(M) - ax
  ages <- nrow(A)
  years <- ncol(A)
  # SVD
  USV <- svd(x = A, nu = ages, nv = years)
  bx <- -USV$u[,1]
  kt <- -USV$d[1]*USV$v[,1]
  # Comment on SVD: Since the svd() function in R orders the mortality
  # indices in increasing order (lowest to highest value), we reverse this by
  # adding a minus sign.

  # Get LC approximated mortality matrix
  Mort <- matrix(nrow = ages, ncol = years)
  for (i in 1:years){
    Mort[,i] <- ax + bx*kt[i]
  }
  # Return list of parameters and matrix of mortality
  return(list(ax, bx, kt, Mort))
}

# Get LC variables and mortality matrix from get.LC.par()
LC <- get.LC(crude.mort)
ax <- unlist(LC[1])
bx <- unlist(LC[2])
kt <- unlist(LC[3])
M.hat <- do.call(rbind, LC[4])
colnames(M.hat) <- 1960:2020
rownames(M.hat) <- 0:90

# Surface plot Lee-carter
persp(seq(0,90)+0.5, seq(1960, 2020), M.hat, theta = -45, phi = 30,
      ticktype = "detailed", xlab = "Age", ylab = "Year",
      zlab = "log fitted mort", col = "lightgrey",
      cex.lab = 0.8, cex.axis = 0.8)

# Plots Lee-Carter variables
par(mfrow = c(2,2), mar = c(4.5,4.5,1,1))
# Plot ax, bx, kt
plot(0:90, ax, xlab = "Age", ylab = expression("a"[x]))
plot(0:90, bx, xlab = "Age", ylab = expression("b"[x]))
plot(1960:2020, kt, xlab = "Year", ylab = expression("k"["t"]))
# Plot Lee-Carter vs data for age=50
plot(1960:2020, log(crude.mort[51,]), xlab="Year", ylab="log(m)")
legend("topright", legend = c("Data", "LC"),
       pch = c(1, -1), lty = c(-1, 1), cex = 0.65)
legend("bottomleft", legend = paste("Age", 50), bty = "n")
lines(1960:2020, M.hat[51,], type="l")

# Create Latex tables of Lee-Carter ax and bx values
library(xtable)
LC.axbx.045 <- cbind(ax[1:46], bx[1:46]) # ages 0-45
colnames(LC.axbx.045) <- c("ax", "bx")
LC.axbx.4690 <- cbind(ax[47:91], bx[47:91]) # ages 46-90
colnames(LC.axbx.4690) <- c("ax", "bx")

tab.axbx.045 <- as.table(LC.axbx.045)
tab.axbx.4690 <- as.table(LC.axbx.4690)
xtable(tab.axbx.045, digits = 6)
xtable(tab.axbx.4690, digits = 6)

# Create Latex table of Lee-Carter kt values
LC.kt.matrix <- cbind(kt)
rownames(LC.kt.matrix) <- 1960:2020
tab.LC.kt <- as.table(LC.kt.matrix)
xtable(tab.LC.kt, digits = 6)

######## Goodness of fit of LC #########

# Calculate ratio of variance explained for LC
years <- 61
ages <- 91
numer <- rep(NA, years)
denom <- rep(NA, years)
eta.sqr <- rep(NA, ages)
for (i in 1:ages){
  for (j in 1:years){
    numer[j] <- (crude.mort[i,j] - exp(M.hat[i,j]))^2
    denom[j] <- (crude.mort[i,j] - exp(ax[i]))^2
  }
  eta.sqr[i] <- 1-(sum(numer)/sum(denom))
}
# Average eta over each age group
eta.sqr.group <- rep(NA, 9) # 9 age groups
eta.sqr.group[1] <- mean(eta.sqr[1:11]) # age group 0-10 yrs
for (i in 1:8){
  eta.sqr.group[i+1] <- mean(eta.sqr[(10*i+2):(10*i+11)])
}
eta.sqr.group

# Plots LC and data for highest and lowest eta
which.min(eta.sqr) # Lowest at index 25 = age 24
which.max(eta.sqr) # Highest at index 81 = age 80

plot(1960:2020, log(crude.mort[25,]), xlab="Year", ylab="log(m)")
legend("topright", legend = c("Data", "LC"),
       pch = c(1, -1), lty = c(-1, 1), cex = 0.65, col=c("black", "blue"))
legend("bottomleft", legend = paste("Age", 24), bty = "n")
lines(1960:2020, M.hat[25,], type="l", col = "blue")

plot(1960:2020, log(crude.mort[81,]), xlab="Year", ylab="log(m)")
legend("topright", legend = c("Data", "LC"),
       pch = c(1, -1), lty = c(-1, 1), cex = 0.65, col=c("black","red"))
legend("bottomleft", legend = paste("Age", 80), bty = "n")
lines(1960:2020, M.hat[81,], type="l", col = "red")


##################################
# Forecasting Norwegian Mortality

# Function for getting forecasted kt and mortality
# kt = previous kt, T = last time point, forc.t = n future years
# seedn = seed number for replicability
get.LC.forc <- function(kt, T, forc.t, seedn, ax, bx){

  # Calculate standard error
  d.hat <- (kt[T] - kt[1])/(T - 1)
  se.sum <- rep(0, length = T-1)
  for (i in 1:T-1){
    se.sum[i] <- (kt[i+1] - kt[i] - d.hat)^2
  }
  se.hat <- sqrt((1/(T - 2))*sum(se.sum))

  # Forecast future kt
  set.seed(seedn)
  kt.hat <- matrix(nrow = forc.t, ncol = 1)
  for (i in 1:forc.t){
    kt.hat[i] <- kt[T] + i*d.hat
    error_term <- rnorm(1, 0, se.hat^2)
    kt.hat[i] <- kt.hat[i] + sqrt(i)*error_term
  }

  # Forcast mortality
  M.forc <- matrix(nrow = 91, ncol = forc.t)
  for (i in 1:forc.t){
    M.forc[,i] <- ax + bx*kt.hat[i]
  }

  # Return forecasted kt and forecasted mort
  return(list(kt.hat, M.forc, se.hat))
}

# Forecasted kt.hat and mortality for 2021-2016 (40 years ahead)
LC.forc <- get.LC.forc(kt, T=years, forc.t=40, seedn=760, ax, bx)
kt.hat <- unlist(LC.forc[1])
M.tilde <- do.call(rbind, LC.forc[2])
se.hat <- unlist(LC.forc[3])
colnames(M.tilde) <- 2021:2060
rownames(M.tilde) <- 0:90

# Plot foreccasted kt
plot(2021:2060, kt.hat, xlab = "Year", ylab = expression("k"[t]),
     main = "Forecasted kt")

# Plot of fitted and forecasted kt
kt.all <- c(kt, kt.hat)
x <- 1960:2060
plot(x, kt.all, type='n', xlab = "Year", ylab = expression("k"[t]),
     main = "Fitted and forecasted kt")
lines(x[x <= 2020], kt.all[1:61], col="green")
lines(x[x >= 2020], kt.all[61:101], col="red")
legend("topright", legend = c("fitted", "forecasted"),
       lty = c(1, 1), col=c("green", "red"), cex = 0.8)

# Forecasted mortality rates surface plot
persp(seq(0,90)+0.5, seq(2021, 2060), M.tilde, theta = -45, phi = 30,
      ticktype = "detailed", zlab= "log(m)",
      xlab = "Age", ylab = "Year", col = "lightgrey",
      main = "Forecasted log mortality", cex.lab = 0.8, cex.axis = 0.8)


# Forecasted mortality rates vs Estimated
# age 0-90 for year 2020, 2030, 2040, 2050 and 2060
color <- rainbow(5)
plot(0:90, M.hat[,61], type="l", col=color[1],
     ylab = "log(m)", xlab = "Age", ylim=c(-12,-2))
legend("bottomright", legend = c("Year 2020", "Year 2030", "Year 2040",
                                  "Year 2050", "Year 2060"),
       cex = 0.7, lty = c(1,1,1,1,1), col = color)
for (i in 1:4){
  lines(0:90, M.tilde[,i*10], col=color[i+1])
}

# Approximated and forecasted mortality rates in one matrix
M.star <- cbind(M.hat, M.tilde)

######## Compare LC forecasted to K2013 #################
years <- 2021:2060
ages <- 0:90
w_male <- 2.671548 - 0.172480*ages +0.001485*(ages^2)
w_female <- 1.287968 - 0.101090*ages + 0.000814*(ages^2)
w_hat <- pmin(rowMeans(cbind(w_male, w_female)), 0)

mort.KOL <- matrix(nrow = 91, ncol = 40)
for (i in 1:40){
  mort.KOL[,i] <- crude.mort[,54]*(1+(w_hat/100))^(2020+i-2013)
}

# Comparison plots
# fixed age plots
df.age <- data.frame(2021:2060, M.tilde[11,], M.tilde[26,], M.tilde[76,],
                     log(mort.KOL[11,]), log(mort.KOL[26,]),
                     log(mort.KOL[76,]))
colnames(df.age) <- c("year", "forc10", "forc25", "forc75",
                      "kol10", "kol25", "kol75")
colors <- c("LC age 10" = "purple", "LC age 25" = "blue",
            "LC age 75" = "green", "K2013 age 10" = "yellow",
            "K2013 age 25" = "orange", "K2013 age 75" = "red")
ggplot(data = df.age, aes(x = age))+
  geom_point(aes(x = year, y = forc10, color="LC age 10")) +
  geom_point(aes(x = year, y = forc25, color="LC age 25"))+
  geom_point(aes(x = year, y = forc75, color="LC age 75"))+
  geom_point(aes(x = year, y = kol10, color="K2013 age 10"), shape=17)+
  geom_point(aes(x = year, y = kol25, color="K2013 age 25"), shape=17)+
  geom_point(aes(x = year, y = kol75, color="K2013 age 75"), shape=17)+
  labs(x = "Year", y = "log(m)", color = "Legend")+
  scale_color_manual(values=colors)+
  scale_shape_manual(name = "Legend",
                     labels = c("LC age 10", "LC age 25", "LC age 75",
                                "K2013 age 10", "K2013 age 25",
                                "K2013 age 75"),
                     values = c(16, 16, 16, 17, 17, 17)) +
  ggtitle("LC forecasts vs K2013")+
  theme(plot.title = element_text(hjust = 0.5))

# fixed year plot for 2060
df.year <- data.frame(0:90, M.tilde[,40], log(mort.KOL[,40]))
colnames(df.year) <- c("age", "LC", "K2013")
colors <- c("LC" = "red", "K2013" = "blue")
ggplot(data = df.year, aes(x = age))+
  geom_point(aes(x = age, y = LC, color="LC")) +
  geom_point(aes(x = age, y = K2013, color="K2013"), shape=18)+
  labs(x = "age", y = "log(m)", color = "Legend")+
  scale_color_manual(values=colors)+
  scale_shape_manual(name = "Legend",
                     labels = c("LC", "K2013")) +
  ggtitle("LC forecasts vs K2013 (year 2060)")+
  theme(plot.title = element_text(hjust = 0.5))

# qx for year 2060
df.year <- data.frame(0:90, 1-(exp(-exp(M.tilde[,40]))),
                               1-exp(-mort.KOL[,40]))
colnames(df.year) <- c("age", "LC", "K2013")
colors <- c("LC" = "red", "K2013" = "blue")
ggplot(data = df.year, aes(x = age))+
  geom_point(aes(x = age, y = LC, color="LC")) +
  geom_point(aes(x = age, y = K2013, color="K2013"), shape=18)+
  labs(x = "age", color = "Legend",
       y=expression(q["x"]))+
  scale_color_manual(values=colors)+
  scale_shape_manual(name = "Legend",
                     labels = c("LC", "K2013")) +
  ggtitle("LC forecasts vs K2013 (year 2060)")+
  theme(plot.title = element_text(hjust = 0.5))


################################################

# Plot approximated and forecasted for years 1960-2020
x <- 1960:2060
na_val <- rep(NA, times=40)
observed <- c(log(crude.mort[51,]), na_val)
plot(x, M.star[51,], type=’n’, xlab = "Year", ylab = "log(m)")
lines(x[x <= 2020], M.star[51,1:61], col="blue")
lines(x[x >= 2020], M.star[51, 61:101], col="red", lty=1)
lines(x[x <= 2020], observed[1:61], type=’p’, col="dark grey")
lines(x[x >= 2020], observed[61:101], type=’p’, col="dark grey")
legend("topright", legend = c("observed", "fitted", "forecasted"),
       pch = c(1, -1, -1), lty = c(-1, 1, 1),
       col=c("dark grey", "blue", "red"), cex = 0.8)
legend("bottomleft", legend = paste("Age", 50), bty = "n")

############# Forecasting accuracy #######################
# Training set from year 1920-1990
D.train <- Deaths[Age <= 90, (Year >= 1920) & (Year <= 1990)]
E.train <- Expos[Age <= 90, (Year >= 1920) & (Year <= 1990)]
mort.train <- get.Crude.M(D.train, E.train)

# Test set from year 1991-2020
D.test <- Deaths[Age <= 90, (Year > 1990) & (Year <= 2020)]
E.test <- Expos[Age <= 90, (Year > 1990) & (Year <= 2020)]
mort.test <- get.Crude.M(D.test, E.test)
```

```
# Lee-Carter on train set
LC.train <- get.LC(mort.train)
ax.train <- unlist(LC.train[1])
bx.train <- unlist(LC.train[2])
kt.train <- unlist(LC.train[3])
M.hat.train <- do.call(rbind, LC.train[4])
colnames(M.hat.train) <- 1920:1990
rownames(M.hat.train) <- 0:90

# Surface plot M.hat.train
persp(seq(0,90)+0.5, seq(1920, 1990), M.hat.train, theta = -45, phi = 30,
      ticktype = "detailed", xlab = "Age", ylab = "Year",
      zlab = "log fitted mort", col = "lightgrey",
      cex.lab = 0.8, cex.axis = 0.8)

# Plot LC train parameters and mortality for age 50
par(mfrow = c(2,2), mar = c(4.5,4.5,1,1))
# Plot ax, bx, kt
plot(0:90, ax.train, xlab = "Age", ylab = expression("a"[x]))
plot(0:90, bx.train, xlab = "Age", ylab = expression("b"[x]))
plot(1920:1990, kt.train, xlab = "Year", ylab = expression("k"["t"]))
# Plot Lee-Carter vs data for age=50
plot(1920:1990, log(mort.train[51,]), xlab="Year", ylab="log(m)")
legend("topright", legend = c("Data", "LC"),
       pch = c(1, -1), lty = c(-1, 1), cex = 0.65)
legend("bottomleft", legend = paste("Age", 50), bty = "n")
lines(1920:1990, M.hat.train[51,], type="l")

# Plot M.hat.train mortality for age 50
plot(1920:1990, log(mort.train[51,]), xlab="Year", ylab="log(m)")
legend("topright", legend = c("Data", "LC"),
       pch = c(1, -1), lty = c(-1, 1), cex = 0.65, col=c("black","red"))
legend("bottomleft", legend = paste("Age", 50), bty = "n")
lines(1920:1990, M.hat.train[51,], type="l", col = "red")


# Lee-Carter forecast years 1991-2020 (30 years)
LC.forc.train <- get.LC.forc(kt=kt.train, T=71,
                             forc.t=30, seedn=880, ax=ax.train,
                             bx=bx.train)
kt.hat.train <- unlist(LC.forc.train[1])
M.tilde.train <- do.call(rbind, LC.forc.train[2])
se.hat.train <- unlist(LC.forc.train[3])
colnames(M.tilde.train) <- 1991:2020
rownames(M.tilde.train) <- 0:90

# Plot estimated and forecasted kt.train
kt.all.train <- c(kt.train, kt.hat.train)
x <- 1920:2020
plot(x, kt.all.train, type='n', xlab = "Year", ylab = expression("k"[t]),
     main = "Fitted and forecasted kt")
points(x[x <= 1990], kt.all.train[1:71], col="green")
points(x[x >= 1990], kt.all.train[71:101], col="red")
legend("topright", legend = c("fitted", "forecasted"),
       pch = c(1, 1), col=c("green", "red"), cex = 0.8)

# LC estimated and forecasted with CI
# Confidence intervals:
CI.upper <- matrix(nrow = 91, ncol = 30)
CI.lower <- matrix(nrow = 91, ncol= 30)
for (i in 1:91) {
  for (j in 1:30){
    CI.lower[i,j] = M.tilde.train[i,j]*
      exp((-1.96)*-bx.train[i]*se.hat.train)
    CI.upper[i,j] = M.tilde.train[i,j]*
      exp(1.96*-bx.train[i]*se.hat.train)
  }
}

# M.hat.train and M.tilde.train in one matrix
M.star.train <- cbind(M.hat.train, M.tilde.train)

# LC forecast plot with confidence intervals
library(tidyverse)
# make dataframe
Age <- 0:90
df.forc <- data.frame(Age, log(mort.test[,30]), M.tilde.train[,30],
                         CI.lower[,30], CI.upper[,30])
colnames(df.forc) <- c("age", "obs", "forc", "CI.low", "CI.up")
# plot
ggplot(data = df.forc, aes(x = age, y = forc)) +
  geom_line(color="red")+
  geom_ribbon(aes(ymin=CI.low, ymax=CI.up), alpha=0.2)+
  geom_point(aes(x = age, y = obs))+
  labs(x = "Age", y = "log(m)")+
  ggtitle("Forecasted vs observed (2020)")+
  theme(plot.title = element_text(hjust = 0.5))


#### Forecast errors #####
e_xt <- mort.test - exp(M.tilde.train)
log_e <- log(mort.test) - M.tilde.train

# plot forecast errors
persp(seq(0:90), 1991:2020, e_xt, theta=45, ticktype="detailed",
      xlab="age", ylab="year", zlab="e", zlim=c(-0.06, 0.01))
title("Plot of forecast errors LC (1991-2020)")

# plot log forecast errors
persp(seq(0:90), 1991:2020, log_e, theta=45, ticktype="detailed",
      xlab="age", ylab="year", zlab="e")
title("Plot of forecast errors LC (1991-2020)")

# Compute MSE
library(Metrics)
MSError <- mse(mort.test, exp(M.tilde.train))
log_MSError <- mse(log(mort.test), M.tilde.train)

MS.df <- data.frame(MSError, log_MSError)
xtable(MS.df, digits=6)

#####################################################
####### Chapter 5: Forecasting reserves ############

# Plot typical trajectory of mortality
stepwise <- function(x){
  return(c(ifelse(x>=55, 1, 0)))
}
age <- 25:115
state <- stepwise(age)

plot(x=age, y=state, type=’s’, axes=FALSE)
axis(1, at = seq(25,115,by=5))
axis(2, at = c(0,1), labels = c("*", expression("\u2020")),
     col = NA, las = 1, pos = 27)


######## Life insurance (endowment) #############
# get transition probabilities from forecasted Lee-Carter (M.tilde)
p_aa <- function(age, year){
  p_surv <- exp(-exp(M.tilde[age+1, year]))
  return(p_surv)
}

# policy functions
a_pol <- function(age){
  a <- rep(0, 2)
  if (age == 66){a[1] = 1000000}
  if (age < 67){a[2] = 2000000}
  return(a)
}

# age at contract start = 30, T = 37, r = 0.03,
# year start of contract = 2021
r = 0.03
res_endow <- rep(0, 38)
res_death <- rep(0, 38)
V_prem <- rep(0, 38)
year_set <- 37:1
# Thiele's difference equation
for (i in 1:37){
  res_endow[38-i] <- exp(-r)*p_aa(67-i, year_set[i])*(a_pol(67-i)[1]
                                                    +res_endow[38-i+1])
  res_death[38-i] <- exp(-r)*(p_aa(67-i, year_set[i])
                            *res_death[38-i+1]+(1-p_aa(67-i, year_set[i]))
                            *(a_pol(67-i)[2]))
  V_prem[38-i] <- -1 + exp(-r)*p_aa(67-i, year_set[i])*V_prem[38-i+1]
}
res_endow[38] <- 1000000
res_total <- res_endow + res_death
pi_0 <- res_total[1]
pi <- -(pi_0/V_prem[1])

# create latex table
ins_mat <- cbind(res_endow, res_death, res_total)
colnames(ins_mat) <- c("PV endowment", "PV death benefit",
                       "PV Total")
rownames(ins_mat) <- 30:67
tab.ins_mat <- as.table(ins_mat)
xtable(tab.ins_mat)

# Calculate present value of premiums
PV_prem <- rep(0, 38)
for (i in 1:37){
  PV_prem[38-i] <- -pi + exp(-r)*p_aa(67-i, year_set[i])*PV_prem[38-i+1]
}

# Create data frame
PV_payout <- res_total
PV_reserve <- PV_payout - PV_prem
PV.df <- data.frame(0:37, PV_prem, PV_payout, PV_reserve)
colnames(PV.df) <- c("age", "premium", "benefits",
                     "reserve")

# plot PV insurance, PV premiums and Mathematical reserves
```

```
colors <- c("PV premium" = "blue", "PV benefits" = "red",
            "Total reserve" = "black")
ggplot(data = PV.df, aes(x = age))+
  geom_point(aes(x =age, y = premium, color="PV premium")) +
  geom_point(aes(x = age, y = benefits, color="PV benefits"))+
  geom_point(aes(x = age, y = reserve, color="Total reserve"))+
  geom_line(aes(x=age, y=0))+
  labs(x = "Age of contract", y = "Present value", color = "Legend")+
  scale_color_manual(values=colors)

# Total reserve for different interest rates r
rates <- c(0.005, 0.015, 0.035, 0.08, 0.20)
res_endow2 <- matrix(0, 38, length(rates))
res_death2 <- matrix(0, 38, length(rates))
year_set <- 37:1
# Thiele's difference equation
for (i in 1:37){
  res_endow2[38-i,] <- exp(-rates)*(p_aa(67-i, year_set[i])*(a_pol(67-i)[1]
                                                      +res_endow2[38-i+1,]))
  res_death2[38-i,] <- exp(-rates)*(p_aa(67-i, year_set[i])
                                   *res_death2[38-i+1,]+
                                     (1-p_aa(67-i, year_set[i]))
                                   *a_pol(67-i)[2])
}
res_endow2[38,] <- 1000000
res_total2 <- res_endow2 + res_death2

# plot Total reserves for the different interest rates
df.rates <- data.frame(0:37, res_total2[,1], res_total2[,2],
                       res_total2[,3], res_total2[,4], res_total2[,5])
colnames(df.rates) <- c("age", "r1", "r2", "r3", "r4", "r5")
colors <- c("r = 0.5%" = "blue", "r = 1.5%" = "green",
            "r = 3.5%" = "yellow", "r = 8%" = "red",
            "r = 20%" = "black")
ggplot(data = df.rates, aes(x = age))+
  geom_point(aes(x = age, y = r1, color="r = 0.5%")) +
  geom_point(aes(x = age, y = r2, color="r = 1.5%"))+
  geom_point(aes(x = age, y = r3, color="r = 3.5%"))+
  geom_point(aes(x = age, y = r4, color="r = 8%"))+
  geom_point(aes(x = age, y = r5, color="r = 20%"))+
  labs(x = "Age of contract", y = "Present value", color = "Legend")+
  scale_color_manual(values=colors)


############ Pension insurance #############
# policy function benefit
a_pre <- function(age){
  a <- 0
  if (age > 66){a = 130000}
  return(a)
}

# premium function
a_tilde <- function(age){
  a <- 0
  if (age < 67){a = 1}
  return(a)
}
# Function for returning yearly survival probabilities pension insurance
p_aa <- function(age, year){
  p_surv <- exp(-exp(M.star[age+1, 41+year]))
  return(p_surv)
}

# Total reserve for different interest rates r
rates <- c(0.01, 0.02, 0.03, 0.08, 0.20)
res_tot <- matrix(0, 61, length(rates))
V_prem <- matrix(0, 61, length(rates))
year_set <- 60:1
# Thiele’s difference equation
for (i in 1:60){
  res_tot[61-i,] <- exp(-rates)*(p_aa(90-i, year_set[i])
                                 *(a_pre(90-i)[1]+res_tot[61-i+1,]))
  V_prem[61-i,] <- -1*a_tilde(90-i) + exp(-rates)*p_aa(90-i, year_set[i])*V_prem[61-i+1,]

}
res_tot[61,] = 0
# calculate pi_0
pi_0 <- res_tot[1,] # for r=3%: 540819.9
pi <- -(pi_0/V_prem[1,]) # for r=3%: 24273.45361

# Caculate present values of premiums
# Calculate present value of premiums
PV_prem <- matrix(0, 61, length(rates))
for (i in 1:60){
  PV_prem[61-i,] <- -pi*a_tilde(90-i) +
    exp(-rates)*p_aa(90-i, year_set[i])*PV_prem[61-i+1,]
}


# Create data frame Present value
PV_payout <- res_tot[,3]
PV_reserve <- PV_payout - PV_prem[,3]
PV.df <- data.frame(0:60, PV_prem[,3], PV_payout, PV_reserve)
colnames(PV.df) <- c("age", "premium", "benefits",
                     "reserve")
# Latex table
ins_mat <- cbind(PV_payout, PV_prem[,3], PV_reserve)
colnames(ins_mat) <- c("PV benefits", "PV premium",
                       "Mathematical reserve")
rownames(ins_mat) <- 30:90
tab.ins_mat <- as.table(ins_mat)
xtable(tab.ins_mat)

# plot PV insurance, PV premiums and Mathematical reserves
colors <- c("PV premium" = "blue", "PV benefits" = "red",
            "Total reserve" = "black")
ggplot(data = PV.df, aes(x = age))+
  geom_point(aes(x = age, y = premium, color="PV premium")) +
  geom_point(aes(x = age, y = benefits, color="PV benefits"))+
  geom_point(aes(x = age, y = reserve, color="Total reserve"))+
  geom_line(aes(x=age, y=0))+
  labs(x = "Age of contract", y = "Present value", color = "Legend")+
  scale_color_manual(values=colors)

# plot PV pension benefits for different interest rates
df.rates <- data.frame(0:60, res_tot[,1], res_tot[,2],
                       res_tot[,3], res_tot[,4], res_tot[,5])
colnames(df.rates) <- c("age", "r1", "r2", "r3", "r4", "r5")
colors <- c("r = 1%" = "blue", "r = 2%" = "green",
            "r = 3%" = "yellow", "r = 8%" = "red",
            "r = 20%" = "black")
ggplot(data = df.rates, aes(x = age))+
  geom_point(aes(x = age, y = r1, color="r = 1%")) +
  geom_point(aes(x = age, y = r2, color="r = 2%"))+
  geom_point(aes(x = age, y = r3, color="r = 3%"))+
  geom_point(aes(x = age, y = r4, color="r = 8%"))+
  geom_point(aes(x = age, y = r5, color="r = 20%"))+
  labs(x = "Age of contract", y = "Present value", color = "Legend")+
  scale_color_manual(values=colors)

# Plot present value of mathematical reserves diff interest rate
# plot PV pension benefits for different interest rates
math_reserve <- res_tot - PV_prem
df.rates <- data.frame(0:60, math_reserve[,1], math_reserve[,2],
                       math_reserve[,3], math_reserve[,4], math_reserve[,5])
colnames(df.rates) <- c("age", "r1", "r2", "r3", "r4", "r5")
colors <- c("r = 1%" = "blue", "r = 2%" = "green",
            "r = 3%" = "yellow", "r = 8%" = "red",
            "r = 20%" = "black")
ggplot(data = df.rates, aes(x = age))+
  geom_point(aes(x = age, y = r1, color="r = 1%")) +
  geom_point(aes(x = age, y = r2, color="r = 2%"))+
  geom_point(aes(x = age, y = r3, color="r = 3%"))+
  geom_point(aes(x = age, y = r4, color="r = 8%"))+
  geom_point(aes(x = age, y = r5, color="r = 20%"))+
  labs(x = "Age of contract", y = "Present value", color = "Legend")+
  scale_color_manual(values=colors)
```

Bibliography

[ALI00] Ahmad, O., Lopez, A., and Inoue, M. "The decline in child mortality: A reappraisal". eng ; por. In: *Bulletin of the World Health Organization* vol. 78, no. 10 (2000), pp. 1175–1191.

[Cai+09] Cairns, A. J. G. et al. "A Quantitative Comparison of Stochastic Mortality Models Using Data From England and Wales and the United States". eng. In: *North American actuarial journal* vol. 13, no. 1 (2009), pp. 1–35.

[DHW19] Dickson, D. C. M., Hardy, M. R., and Waters, H. R. *Actuarial Mathematics for Life Contingent Risks*. 3rd ed. International Series on Actuarial Science. Cambridge University Press, 2019.

[GK08] Girosi, F. and King, G. *Demographic Forecasting*. Princeton: Princeton University Press, 2008.

[Gom25] Gompertz, B. "On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies". In: *Philosophical Transactions of the Royal Society of London* vol. 115 (1825), pp. 513–583.

[HA21] Hyndeman, R. J. and Athanasopoulos, G. *Forecasting: principles and practice 3rd, edition*. Ed. by OTexts: Melbourne, A. 2021. URL: https://otexts.com/fpp3/ (visited on 04/28/2022).

[KAM16] Khan, M. H. R., Afrin, S., and Masud, M. S. "Mortality Forecasting Using Lee Carter Model Implemented to French Mortality Data". eng. In: *The Dhaka University journal of science* vol. 64, no. 2 (2016), pp. 99–104.

[Kol] Koller, M. *Stochastic Models in Life Insurance*. eng. 2012th ed. EAA Series. Berlin, Heidelberg: Springer Berlin Heidelberg.

[LC92] Lee, R. D. and Carter, L. R. "Modeling and Forecasting U.S. Mortality". In: *Journal of the American Statistical Association* vol. 87, no. 419 (1992), pp. 659–671. eprint: https://doi.org/10.1080/01621459.1992.10475265.

[MRC18] Macdonald, A. S., Richards, S. J., and Currie, I. D. *Modelling Mortality with Actuarial Applications*. International Series on Actuarial Science. Cambridge University Press, 2018.

[R C22] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2022.

[Sad12] Sadek, R. “SVD Based Image Processing Applications: State of The Art, Contributions and Research Challenges”. In: *International Journal of Advanced Computer Science and Applications - IJACSA* vol. 3 (Nov. 2012).

[Sma94] Small, C. “A global analysis of mid-ocean ridge axial topography”. In: *Geophysical Journal International* vol. 116, no. 1 (Jan. 1994), pp. 64–84. eprint: https://academic.oup.com/gji/article-pdf/116/1/64/1790410/116-1-64.pdf.

[TB97] Trefethen, L. N. and Bau III, D. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics (SIAM), 1997.

[WRR02] Wall, M., Rechtsteiner, A., and Rocha, L. “Singular Value Decomposition and Principal Component Analysis”. In: *In A Practical Approach to Microarray Data Analysis* vol. 5 (Sept. 2002).