Application of a power-exponential function based model to mortality rates forecasting

by

Andromachi Boulougari

MASTER THESIS IN MATHEMATICS/ APPLIED MATHEMATICS
DIVISION OF APPLIED MATHEMATICS
MÅLARDALEN UNIVERSITY
SE-721 23 VÄSTERÅS, SWEDEN
Master thesis in mathematics / applied mathematics

Date:
2018-05-31

Project name:
Application of a power-exponential function based model to mortality rates forecasting

Author:
Andromachi Boulougari

Supervisor(s):
Milica Rančić and Karl Lundengård

Reviewer:
Anatoliy Malyarenko

Examiner:
Ying Ni

Comprising:
30 ECTS credits
Dedicated to my heroes, mom and dad.
Acknowledgements

I would like to express my thanks and gratitude to my supervisor Milica Rančić for her supportive and positive guidance as well as to my supervisor Karl Lundengård for his help in the use of MATLAB. Special thanks to the reviewer, Professor Anatoliy Malyarenko, for giving a very detailed feedback and to the examiner, Dr. Ying Ni, for this thesis.

Last but not least, I want to thank my family and friends for their support and encouragement.
Abstract

The modeling of a law of mortality has been a consistent interest from a vast majority of researchers and many models through the years have been suggested. The first aim of this thesis is to systematically evaluate a selection of models — Modified Perks, Heligman-Pollard and Power-exponential — to determine their relative strengths and weaknesses with respect to forecasting the mortality rate using Lee-Carter model. The second aim is to fit death rates data by the selective models from USA, Sweden and Greece using numerical techniques for curve-fitting with the non-linear least squares method. The results indicate that the Heligman-Pollard model performs better especially when the phenomenon of the “accident hump” occurs during adulthood.

**Keywords:** Mortality rates, Modified Perks model, Heligman–Pollard model, Power-exponential model, Forecasting, Extrapolation, Lee-Carter, Non-linear least squares method, Curve fitting.
# Contents

1 Introduction .................................................. 6
   1.1 Mortality Rate Models .................................... 7
   1.2 Methods for Mortality Forecasting ...................... 10

2 Selected Mortality Models ................................. 11
   2.1 The Modified Perks Model ................................. 12
   2.2 The Heligman–Pollard Model — HP4 ....................... 16
   2.3 Power-exponential Function Based Model ................. 24

3 Selected Methods for Forecasting ....................... 29
   3.1 The Lee–Carter Model ..................................... 29
   3.2 Regression and the Least Square Method ................. 31
      3.2.1 Linear Regression .................................. 31
      3.2.2 Non–Linear Regression .............................. 33

4 Model Fitting ................................................. 34
   4.1 Analysis of Mortality Data ................................. 34
   4.2 Forecasting Mortality Rates .............................. 38

5 Conclusion .................................................. 41
Summary of reflection of objectives in the thesis

6.1 Objective 1: Knowledge and understanding

6.2 Objective 2: Methodological knowledge

6.3 Objective 3: Critically and Systematically Integrate Knowledge

6.4 Objective 4: Independently and Creatively Identify and Carry out Advanced Tasks

6.5 Objective 5: Present and Discuss Conclusions and Knowledge

6.6 Objective 6: Scientific, Social and Ethical Aspects
List of Figures

2.1 Example of the mortality rate curve given by the Modified Perks Model in the year 2010. 13

2.2 The effect on the logarithm of the mortality rate curve when one parameter is varied, while the others retain the values given in Table 2.1. 14

2.3 Example of the mortality rate curve given by the HP4 Model in the year 2010. 18

2.4 The effect on the logarithm of the mortality rate curve when “Term 1” is varied, while the others retain the values given in Table 2.2. 18

2.5 The effect on the logarithm of the mortality rate curve when “Term 2” is varied, while the others retain the values given in Table 2.2. 19

2.6 The effect on the logarithm of the mortality rate curve when “Term 3” is varied, while the others retain the values given in Table 2.2. 19

2.7 Example of the mortality rate curve given by the Power-exponential Model in the year 2010. 25

2.8 The effect on the logarithm of the mortality rate curve when one parameter is varied, while the others retain the values given in Table 2.3. 26

4.1 Mortality rates fitted by the three models for male population of USA. Figure 4.1a is the population for the year 1970, Figure 4.1b is population for the year 2000 and Figure 4.1c is population for the year 2010. 35

4.2 Mortality rates fitted by the three models for male population of Sweden. Figure 4.2a is population for the year 1970, Figure 4.2b is population for the year 2000 and Figure 4.2c is population for the year 2010. 36

4.3 Mortality rates fitted by the three models for male population of Greece. Fig. 4.3a is population for the year 1981, Fig. 4.3b is population for the year 2000 and Fig. 4.3c is population for the year 2010. 37
4.4 Forecasting mortality rates for male population of USA in the years 2000–
2010. The figure displays the forecasting for the last year. .......................... 39

4.5 Forecasting mortality rates for male population of Sweden in the years 2000–
2010. The figure displays the forecasting for the last year. .......................... 39

4.6 Forecasting mortality rates for male population of Greece in the years 2001–
2011. The figure displays the forecasting for the last year. .......................... 40
Chapter 1

Introduction

When it comes to the topic of risk, most of us will readily agree that risk has become a focal point for anyone involved in financial markets. Indeed, the financial market crisis that began in 2007 confirmed the centrality of risk analysis in finance and extended beyond the economic sector in the United States to the entire world. In this project, we focus on the risk involved in insurance companies and more specifically the financial impact of the unexpected mortality in that sector. For that reason, over the last decades, the modeling of a law of mortality has been a consistent interest from a vast majority of researchers [11, 18]. The goal of this thesis is to systematically evaluate a few different models to determine their relative strengths and weaknesses with respect to forecasting the mortality rate.

In the following sections of the introduction, a brief review of existing models for mortality rates is described along with different methods of forecasting. As will be seen in Chapter 2, after a selection of papers and research on mortality rates, we will discuss about the Modified Perks model [4], the Heligman–Pollard (HP4) model [11] and the Power-exponential one proposed by Lundengård et al [16] and we will derive survival functions corresponding to all three models. In Chapter 3, we will discuss which method we intend to use for forecasting and how to measure the quality of the forecast. In Chapter 4, several different models will be fitted to data for measured death rates from many different countries using numerical techniques for curve-fitting with the non-linear least squares method. Then, the resulting mortality rates will be used to examine suitable methods for forecasting the mortality rates for a given model. Part of the results obtained in this thesis will be reported in the full version of [15]¹. Finally, conclusions will follow in Chapter 5.

¹The full version of [15] will be presented at the 5th International Conference on Stochastic Modeling Techniques and Data Analysis, SMTDA 2018, in June 2018.
1.1 Mortality Rate Models

In fact, there is a wide range of models that describes the death rate of an individual. The following models follow a typical mortality pattern which distributes mortality into three key stages: 1) infant mortality — mortality decreases with age — 2) “accident hump” — mortality increases due to accidents — and 3) senescent mortality — mortality follows the Gompertz law [7, 9]. The first suggestion for the law of mortality as a mathematical expression has been made by Benjamin Gompertz in 1825. In his study, he explains how the force of mortality and age are linked — mortality rates exponentially increase with age [8]. The mathematical expression of this simple mortality law is the following:

\[ \mu(x) = b \exp(cx) , \]

where \( \mu \) is the force of mortality of an individual aged \( x \); \( b \) is the initial mortality at age 0; and \( c \) is the mortality’s rate (geometric increase) of change as the age increases. Even though this law can be successfully used to model mortality rates in different populations and ages, many researchers in the field have modified the model to smoothly justify deviations from age to age as well as accurately display the pattern that mortality follows[7, 11]. In what follows, we will discuss an extension of the Gompertz law which is called Gompertz–Makeham law and its mathematical expression is:

\[ \mu(x) = a + b \exp(cx) . \]

This adding parameter, \( a \), is an associated risk that gives insight into causes of mortality that the first law could not capture [21]. The parameter \( b \) measures the mortality rate in younger ages. It is important to highlight that even though the particular model performs better in earlier ages, it makes a gradual increase when it reaches the older ages where the Gompertz law is performed [22]. Another extension of the two mortality models is the so-called Double Geometric and has five parameters:

\[ \mu(x) = a + b_1 b_2^x + c_1 c_2^x . \]

Later, Thiele suggested a model that includes seven parameters which successfully display the three aforementioned periods of mortality in an individual’s life span: infant mortality, accident hump, and senescent mortality [25]. Respectively, the model’s form is:

\[ \mu(x) = a_1 \exp(-b_1 x) + a_2 \exp\left(-b_2 \frac{(x-c)^2}{2}\right) + a_3 \exp(b_3 x) . \]

Furthermore, there are several frailty models which are modifications of the Gompertz law. Those models use frailty distribution to measure heterogeneity in mortality data. An extensive explanation is given in Chapter 2 from [4]. Perks proposed two models that follow a gamma distribution: 1) the Gompertz-gamma, known as Perks and 2) the Gompertz/Makeham-gamma, known as Modified Perks. The difference of these two models from the initial law is that in
the Gompertz model there is no presence of heterogeneity — any variations in future survival that is possibly observable at the beginning of an study [4]. The mathematical expression of the Perks and Modified Perks is respectively as follows:

\[
\mu(x) = \frac{a}{1 + \exp(b - cx)}, \\
\mu(x) = \frac{a}{1 + \exp(b - cx)} + d.
\]

This model is more preferable to measure mortality rates because it uses gamma frailty distribution which is tractable, flexible as well as has positive frailty [26]. In fact, this is true because a more recent study from Butt and Haberman [4] has shown that the use of the gamma frailty distribution provides better fit during the whole lifespan along with a monotonous decrease of heterogeneity in older ages due to the fact that the model is following a logistic form with age; meaning that the graph of mortality approaches \( a \) as \( x \) approaches \( +\infty \) [4, 7, 14]. Another frailty model that incorporates Gompertz law is the Gompertz-inverse Gaussian. Its formula is:

\[
\mu(x) = \frac{\exp(a - bx)}{\sqrt{1 + \exp(-c + bx)}}.
\]

However, this model is not in favor because, based on the aforementioned study, it cannot show satisfactory results regarding the rates of mortality in older ages [4].

Also, Weibull proposed a model that due to its flexibility depending on the two parameters can be applied not only to technical devices but also to human mortality:

\[
\mu(x) = \frac{a}{b} \left( \frac{x}{b} \right)^{a-1},
\]

where the probability of surviving depends on the probability of someone not dying from various causes [22, 27].

Following, a model that incorporates the Gompertz law is the one proposed by Heligman and Pollard (HP). The eight parameters model has four similar variations and all strive to cover the whole lifespan. The initial model HP1 is as follows:

\[
\mu(x) = a_1^{(x+a_2)^{y_3}} + b_1 \exp \left( -b_2 \ln \left( \frac{x}{b_3} \right)^2 \right) + c_1 c_2^x.
\]

The HP2 is similar to HP1 because the values of the first two terms are very small in older ages [7]. Thus, HP2 formula looks like:

\[
\mu(x) = a_1^{(x+a_2)^{y_3}} + b_1 \exp \left( -b_2 \ln \left( \frac{x}{b_3} \right)^2 \right) + \frac{c_1 c_2^x}{1 + c_1 c_2^x}.
\]
Furthermore, the following two variations include nine parameters improving model fit:

\[ \begin{align*}
    \text{HP3} : \mu(x) &= a_1^{(x+a_2)^{a_3}} + b_1 \exp \left( -b_2 \ln \left( \frac{x}{b_3} \right)^2 \right) + \frac{c_1 c_2^x}{1 + c_3 c_1^c} , \\
    \text{HP4} : \mu(x) &= a_1^{(x+a_2)^{a_3}} + b_1 \exp \left( -b_2 \ln \left( \frac{x}{b_3} \right)^2 \right) + \frac{c_1 c_2^{c^x}}{1 + c_1 c_2^{c^x}} .
\end{align*} \]

In all four variations, each term corresponds to each stage of mortality, where in the last part the Gompertz law is included [11]. The HP4 will be extensively explained in Chapter 2.

More recently, a logistic model is suggested that includes three parameters:

\[ \mu(x) = \frac{a \exp(bx)}{1 + \frac{a c}{b} \left( \exp(bx) - 1 \right)} , \]

which predicts higher ages better in a specific study [22].

Another common model is the log-logistic which provides a two-parametric model for mortality analysis:

\[ \mu(x) = \frac{abx^{a-1}}{1 + bx^a} . \]

Finally, Hannerz proposed a five-parameter formula:

\[ \mu(x) = \frac{g(x) \exp G(x)}{1 + \exp G(x)} , \]

where

\[ g(x) = \frac{a_1}{x^2} + a_2 x + a_3 \exp(cx) , \]

and

\[ G(x) = a_0 - \frac{a_1}{x} + \frac{a_2 x^2}{2} + \frac{a_3}{c} \exp(cx) , \]

where \( a_i, i = 0, 1, 2, 3 \), are the parameters that denote population’s mortality differences and \( c \) is a constant that relates different mortality schedules [10].

Last but not least, the power-exponential function based model proposed by Lundengård et al [16] includes five parameters and its formula is:

\[ \mu(x) = \frac{c_1}{x e^{-c_2 x}} + a_1 \left( x e^{-a_2 x} \right)^{a_3} . \]

The parameters \( c_1, c_2, a_1, a_2, a_3 \) can easily be interpreted in terms of qualitative properties of the curve [16]. The model will be extensively explained in Chapter 2.

Following, we will give a brief review of the different methods of forecasting.
1.2 Methods for Mortality Forecasting

As previously mentioned, many researchers for a long period of time have highlighted the importance of mortality modelling. However, recently, due to rapid aging of the population, a more important issue has been brought to the fore; forecasting mortality rates. The majority of those involved in this field will readily agree that the methods for population forecasting were more simple and based on subjectivity in the past decades, while the new ones are more sophisticated and complex but promising [3, 24]. The methods for forecasting are distributed into three main approaches: 1) explanation — forecasting based on structural or epidemiological models — 2) expectation — forecasting based on expert opinion — and 3) extrapolation — forecasting based on patterns and trends [2, 24].

Firstly, the less used approach is the explanatory methods. This methods are mainly based on structural models and used for short-term forecasting. The main advantages of those models are that their feedback mechanisms and limiting factors can be taken into account as well as they are assumed as the ideal method for forecasting mortality. Regardless those advantages, they have not produced reliable results yet, due to the lack of independence between causes of death and insufficient data [2, 3]. Another promising approach for short-term forecasting is the so-called expectation. This approach involves three different kinds of expectations: i) individual, ii) group of experts, and iii) informed judgment of experts. However, such methods have tended to fail in the forecasting process due to subjectivity as well as lag prevailing trends [2].

Furthermore, the most common method for forecasting mortality rates is the extrapolation. In fact, many developments have been made in extrapolative methods. Particularly, those methods involve models than can be classified from zero to three factors models. Those factors can be either age, period or cohort [3]. Depending on the number of factors, those models can be either an ARIMA model (zero-factor model), parametrisation — known as laws of mortality — (one-factor model), Lee–Carter model and its extensions or Generalised Linear Modelling/GLM (two-factors model), and finally cohort models or APC (three-factors model). Among all the aforementioned methods, the parametrisation — specifically the use of HP4 model — and the Lee–Carter model are the most applicable and successful ones for long-term mortality forecasting in many developed countries [2, 3, 24]. Even though parametrisation is useful for forecasting, many difficulties are rising in terms of determining the best period in order to project the parameters. For that reason, Lee–Carter model is more preferable and we will use it in our forecasting process, see also Chapter 3.
Chapter 2

Selected Mortality Models

In this Chapter, we will focus on three mortality models; namely, Modified Perks, Heligman–Pollard (HP4) and Power-exponential function based model. Firstly, we will provide the reader with the required theory based on the force of mortality. Following, we will derive the survival functions corresponding to all three models.

Definition 1. We define the force of mortality as:

\[ \mu(x) = \lim_{dx \to 0^+} \frac{1}{dx} \Pr[T_x \leq dx] \]  \hspace{1cm} (2.1)

The death of an individual aged \( x \) can happen at any age that is greater than \( x \). We model the future lifetime of an individual by a continuous random variable which we denote by \( T_x \).

See also [14, Definition 2.2] for a formal definition of a continuous random variable. Therefore, \( x + T_x \) represents the age-at-death random variable for \( x \).

Theorem 1 (Lifetime distribution). We denote \( F_x \), the probability that someone does not survive beyond age \( x + t \), known as lifetime distribution:

\[ F_x(t) = \Pr[T_x \leq t] \]  \hspace{1cm} (2.2)

Theorem 2 (Survival function). We assume that the force of mortality is constant between integer ages. Thus, for integer \( x \) and \( 0 < s < 1 \), we assume that \( \mu(x + s) \) does not depend on the time \( s \). The survival function \( S_x \) is defined by:

\[ S_x(t) = 1 - F_x(t) = \Pr[T_x > t] = \exp \left( - \int_0^t \mu(x + s) \, ds \right) \]  \hspace{1cm} (2.3)

However, in order to get proper survival function for a lifetime distribution, two conditions must be satisfied: 1) the probability that someone currently aged \( x \) survives 0 years is equal to 1 and 2) all eventually die. Respectively, the conditions in mathematical expressions are:

\[ S_x(0) = 1 \]  \hspace{1cm} (2.4)
\[
\lim_{t \to \infty} S_x(t) = 0 .
\] (2.5)

The proof of the Theorem 1 and 2 along with the properties (2.4) and (2.5) can be found in the book [5].

### 2.1 The Modified Perks Model

The first model that we will take into consideration is the Modified Perks. As previously mentioned in Introduction, this model is more preferable to measure mortality rates because it uses gamma frailty distribution which is tractable, flexible as well as has positive frailty [26]. Frailty is defined as an unobservable statistical variable \( Z \) that describes implications for the standard life table methods [4]. The reader can refer to the two aforementioned sources for further explanations.

**Definition 2** (Modified Perks Model – MP). The mathematical expression of the Modified Perks mortality rate model is defined as

\[
\mu(x) = \frac{a}{1 + \exp(b - cx)} + d .
\] (2.6)

We gathered data regarding the Swedish population from the Human Mortality Database and we fitted the Modified Perk model to this data to estimate the values for the four parameters for many different years [12]. The parameter estimates for the year 2010 are given in the following table as well as a graphic representation of the resulting mortality rate curve of the model is given in Fig. 2.1.

As seen in Eq. (2.6), the model contains four parameters that modify the shape of the mortality curve for the Modified Perks model. Following, we provide different figures that show the effect on the mortality rate curve when a unique parameter is varied, while the other parameters retain the values given in Table 2.1. It is important to highlight that the “Term 1” represents the mortality rate in the higher ages while the “Term 2” the younger ones.

From Fig. 2.2 we can observe which parameters are sufficient to render the different shapes that the logarithm of the mortality rate curve can take. More specifically, \( a \) — representing the force of mortality — and \( c \) — representing the relative rate of increase in \( a \) with the age [4].

<table>
<thead>
<tr>
<th>( a )</th>
<th>( b )</th>
<th>( c )</th>
<th>( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>805651.726</td>
<td>24.568</td>
<td>0.101</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Figure 2.1: Example of the mortality rate curve given by the Modified Perks Model in the year 2010.
Figure 2.2: The effect on the logarithm of the mortality rate curve when one parameter is varied, while the others retain the values given in Table 2.1.
Therefore, a decrease in the parameter $a$ results with a shift of the curve to the right while the graph increases faster as parameter $c$ increases resulting with a less flatter curve (Figs. 2.2a, 2.2c). In Fig. 2.2c, we do not get any mortality rates in the elder ages implying that death is more likely to occur in the earlier ages, even though this is not realistic. Thus, in general, as “Term 1” increases we get a non-flat curve with a fast force of mortality. On the other side, the parameter $d$ (red line) shows us the logarithm of the mortality rate in younger ages 0–10 while term 1 remains the same, see also Fig. 2.2d. We can observe that the curve moves upwards or downwards in a positive association with the values of $d$. However, even though the Modified Perks model seems to fit the curve due to their logistic form with respect to age, it fails to show the “accident hump” observed in the mortality rate during adulthood. For that reason, other mortality rate models have been proposed with more parameters that provide the users with a better fit.

**Theorem 3** (Modified Perks Model – Survival Function). *Given the Theorem 2 we get that the survival function of Modified Perks is*

$$S_{MP}^x(t) = \exp \left\{-\frac{a}{c} \left[ \ln \left( \exp \left( c(x+t) \right) + \exp(b) \right) - \ln \left( \exp(cx) + \exp(b) \right) \right] - dt \right\}.$$  \hspace{1cm} (2.7)

*The survival function satisfies the two conditions:*

1. $S_{MP}^x(0) = 1$,
2. $\lim_{t \to \infty} S_{MP}^x(t) = 0$.

**Proof.** For computational convenience we will calculate $\ln S_{MP}^x(t)$

$$-\ln S_{MP}^x(t) = \int_0^t \frac{a}{1 + \exp \left( b - c(x+s) \right)} ds + \int_0^t \frac{d}{ds} ds$$

and separate the above equation in two parts:

i) $$\int_0^t \frac{a}{1 + \exp \left( b - c(x+s) \right)} ds = a \exp(cx) \int_0^t \frac{\exp(cs)}{\exp(c(x+s)) + \exp(b)} ds.$$  

We consider the change of variable $u = \exp(c(x+s)) + \exp(b)$ so that $du = c \exp(c(x+s)) ds$. Therefore, we get:
$$a \exp(cx) \int \frac{1}{\exp(cx) u} \, du = \frac{a}{c} \int \frac{1}{u} \, du$$

$$= \frac{a}{c} \ln \left( \frac{\exp(c(x+t)) + \exp(b)}{\exp(cx) + \exp(b)} \right)$$

$$= \frac{a}{c} \left[ \ln \left( \exp(c(x+t)) + \exp(b) \right) - \ln \left( \exp(cx) + \exp(b) \right) \right].$$

\( (2.8) \)

\[ ii) \]

\[ \int_0^t d \tau = \int_0^t d \tau = d S \bigg|_0^t = d t \ . \quad (2.9) \]

Combining both Eq. (2.8) and Eq. (2.9) we get Eq. (2.7).

Since \( S_{MP}^x(t) \) is continuous and the two conditions are satisfied:

\[ S_{MP}^x(0) = \exp \left\{ -\frac{a}{c} \left[ \ln \left( \exp(cx) + \exp(b) \right) - \ln \left( \exp(cx) + \exp(b) \right) \right] \right\} = e^0 = 1 , \]

\[ \lim_{t \to \infty} S_{MP}^x(t) = \lim_{t \to \infty} \exp \left\{ -\frac{a}{c} \left[ \ln \left( \exp(c(x+t)) + \exp(b) \right) - \ln \left( \exp(cx) + \exp(b) \right) \right] - dt \right\} = 0 , \]

where \( \lim_{t \to \infty} (e^{c(x+t)} + e^b)^{-\frac{a}{c}} = \infty \) and \( \lim_{t \to \infty} t^{-\frac{a}{c}} = 0 \), both \( S_{MP}^x(t) \) and \( F_{MP}^x(t) \) are proper cumulative distribution functions.

\[ 2.2 \quad \text{The Heligman–Pollard Model — HP4} \]

A continuous interest has been risen around another mortality law known as Heligman–Pollard (HP1) that can adequately represent the pattern of the force of mortality. The model contains three terms, each representing the three main periods of mortality in an individual’s life span: infant mortality, accident hump, and senescent mortality [25]. In this section, we focus on an extension of the initial model known as HP4. Due to the difficulty to fit the real “accident hump” that was at the younger ages close to 20, Heligman and Pollard proposed HP4 which key advantage is that it performs better and fits better at older ages comparing to the rest of extensions [11].
Definition 3 (Heligman Pollard Model — HP4). The mathematical expression of the mortality rate model is defined as

\[ \mu(x) = a_1 x^{a_2 x^{a_3}} + b_1 \exp(-b_2 \ln\left(\frac{x}{b_3}\right)^2) + \frac{c_1 c_2^{c_3}}{1 + c_1 c_2^{c_3}}. \]  

(2.10)

According to the initial paper [11]:

- \(a_1\) is the level of infant mortality
- \(a_2\) is the location of infant mortality
- \(a_3\) is the rate of infant mortality decline
- \(b_1\) is the severity of “accident hump”
- \(b_2\) is the spread of “accident hump”
- \(b_3\) is the location of “accident hump”
- \(c_1\) is the base level of senescent mortality
- \(c_2\) is the rate of senescent mortality increase
- \(c_3\) is the curvature

Likewise, we gathered data and we fitted the HP4 model to the data to estimate the values for the nine parameters for many different years. The parameter estimates for the year 2010 are given in Table 2.2 as well as a graphic representation of the resulting mortality rate curve of the model is given in Figure 2.3.

As seen in Eq. (2.10), the model contains various parameters that modify the shape of the mortality curve for the Modified Perks model. Following, we provide different figures that show the effect on the mortality rate curve when a unique parameter is varied, while the other parameters retain the values given in Table 2.2. As previously mentioned, “Term 1” represents the infant mortality, “Term 2” the accident hump, and “Term 3” the senescent mortality.

First, from Fig. 2.4 we can observe that the parameters that have a significant effect on the initial curve are \(a_2\) and \(a_3\). Thus, the parameter \(a_1\) does not affect the curve as much as the other

Table 2.2: Parameter estimates for the logarithm of the mortality rate fitted to mortality in the Swedish population 2010.

<table>
<thead>
<tr>
<th>(a_1)</th>
<th>(a_2)</th>
<th>(a_3)</th>
<th>(b_1)</th>
<th>(b_2)</th>
<th>(b_3)</th>
<th>(c_1)</th>
<th>(c_2)</th>
<th>(c_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9999</td>
<td>8.4773</td>
<td>8.4526</td>
<td>0.0004</td>
<td>15.4891</td>
<td>23.347</td>
<td>0.0001</td>
<td>1.014</td>
<td>1.4086</td>
</tr>
</tbody>
</table>
Figure 2.3: Example of the mortality rate curve given by the HP4 Model in the year 2010.

Figure 2.4: The effect on the logarithm of the mortality rate curve when “Term 1” is varied, while the others retain the values given in Table 2.2.
Figure 2.5: The effect on the logarithm of the mortality rate curve when “Term 2” is varied, while the others retain the values given in Table 2.2.

Figure 2.6: The effect on the logarithm of the mortality rate curve when “Term 3” is varied, while the others retain the values given in Table 2.2.
two. However, by decreasing the parameter $a_2$ the yellow vertical line representing “Term 1” just moves to the right and then follows the initial curve, while a decrease in parameter $a_3$ results with a bigger movement to the right (see pink vertical line) but with a rapidly declining exponential reflecting the fall in the level of mortality during the early childhood years. Then, by equally decreasing both parameters $a_2$ and $a_3$, we get that the level of mortality starts at the age 20 where the actual “accident hump” occurs having no data regarding individual’s first years of life.

The “Term 2” reflects the accident mortality for the population. The “accident hump” is found in all populations and appears as a distinct hump in the mortality curve generally between the ages 10 and 40 [11]. First of all, changes in the parameter $b_1$ have an effect on the hump height in the mortality curve. The curve of “Term 2” can be more narrow or wide resulting respectively with a lower or higher hump, see also Fig. 2.5. For example, in the case where parameter $b_2$ decreases we get a wider spread resulting in an increase in the range of ages when the distinct hump in the mortality curve appears. Finally, a change in the parameter $b_3$ reflects on the location of the “accident hump” and the age range when the hump can occur.

Finally, the last term represents the senescent mortality and includes the well-known law of mortality, the Gompertz law. In Fig. 2.6 we can conclude that there is a geometric increase in the mortality rate in the older ages. More specifically, increased $c_1$ results with a straight line. In addition, an increase in the values of $c_2$ and $c_3$ results with an increased curvature downward (flatter curve with no hump), whereas a decrease in those parameters denote an increased curvature upward (steeper curve with bigger hump).

**Theorem 4** (HP4 – Survival Function). *Given the Theorem 2 we get that the survival function of Heligman Pollard Model – HP4 is*

\[
S_{x}^{HP4}(t) = \exp \left\{ - \frac{1}{a_3} \left( \frac{-1}{\ln a_1} \right) \left( \gamma \left( \frac{1}{a_3}, -\ln a_1(x + t + a_2)^{a_3} \right) \right) - \gamma \left( \frac{1}{a_3}, -\ln a_1(x + a_2)^{a_3} \right) \right\}
\]

\[
- b_1 b_3^{2b_2} \left[ \frac{(x + t)^{-2b_2 + 1}}{-2b_2 + 1} - \frac{x^{-2b_2 + 1}}{-2b_2 + 1} \right]
\]

\[
- \int_{0}^{t} \frac{c_1 c_2^{(x+s)^{3}}}{1 + c_1 c_2^{(x+s)^{3}}} ds \right\},
\]

where \(\gamma(a, t) = \int_{0}^{t} x^{a-1} e^{-x} dx\) is the lower incomplete Gamma function [1].

The survival function satisfies the two conditions:

1. \(S_{x}^{HP4}(0) = 1\),

2. \(\lim_{t \to \infty} S_{x}^{HP4}(t) = 0\).
Proof. Once again, for computational convenience we will calculate $\ln S^H_{\mathcal{P}}(t)$ and each part separately:

i) We set $u = (x + s + a_2)^{a_3}$ so that $du = a_3(x + s + a_2)^{a_3-1} \; ds$. Following, we set $w = -u \ln a_1$ where $dw = -\ln a_1 \; du$; thus, we obtain

\[
\int_0^t \left( a_1(x+s+a_2)^{a_3} \right) \; ds = \int_0^t \left( x + t + a_2 \right)^{a_3} \left( \frac{a_1^u}{a_3 u^{1 - a_3}} \right) \; du \\
= \frac{1}{a_3} \int_0^t (x + t + a_2)^{a_3} e^{\ln (a_1) u \frac{1}{a_3} - \frac{1}{a_3} \ln a_1} \; du \\
= \frac{1}{a_3} \int_0^t (x + t + a_2)^{a_3} e^{w \frac{1}{a_3} - \frac{1}{a_3} \ln a_1} \; dw \\
= \frac{1}{a_3} \left( \frac{-1}{\ln a_1} \right) \frac{1}{a_3} \left[ \ln a_1 (x + t + a_2)^a \right] \frac{-1}{a_3} \ln a_1 (x + t + a_2)^a \\
= \frac{1}{a_3} \left( \frac{-1}{\ln a_1} \right) \left[ \gamma\left( \frac{1}{a_3}, w \right) \right] \frac{-1}{a_3} \ln a_1 (x + t + a_2)^a \\
= \frac{1}{a_3} \left( \frac{-1}{\ln a_1} \right) \left[ \gamma\left( \frac{1}{a_3}, \ln a_1 (x + t + a_2)^a \right) \right. \\
\left. \left. - \gamma\left( \frac{1}{a_3}, \ln a_1 (x + a_2)^a \right) \right) \right]. \quad (2.12)
\]

ii) We consider the change of variable $u = x + s$ so that $du = ds$. Therefore, we get:

\[
\int_0^t b_1 \exp \left( -b_2 \ln \left( \frac{x+s}{b_3} \right)^2 \right) \; ds = b_1 \int_0^t \exp \left( \ln \left( \frac{b_3}{x+s} \right)^{2b_2} \right) \; ds \\
= b_1 b_3^{2b_2} \int_0^t \frac{1}{(x+s)^{2b_2}} \; ds \\
= b_1 b_3^{2b_2} \int_x^{x+t} \frac{1}{u^{2b_2}} \; du \\
= b_1 b_3^{2b_2} \left[ \frac{u^{-2b_2+1}}{-2b_2+1} \right]_x^{x+t} \\
= b_1 b_3^{2b_2} \left( (x+t)^{-2b_2+1} - x^{-2b_2+1} \right). \quad (2.13)
\]
iii) In this part, we leave the “Term 3” as it is because it must be numerically solved.

\[\int_{0}^{t} \frac{c_1c_2^{(x+s)^{\gamma_3}}}{1 + c_1c_2^{(x+s)^{\gamma_3}}} \, ds. \quad (2.14)\]

Combining Eq. (2.12), Eq. (2.13) and Eq. (2.14) we get Eq. (2.11).

\[S_{x}^{HP4}(t)\] is continuous and the two conditions are satisfied:

\[S_{x}^{HP4}(0) = \exp \left\{ -\frac{1}{a_3} \left( \frac{-1}{\ln a_1} \right)^{\frac{1}{\gamma_3}} \left( \frac{1}{a_3}, -\ln a_1(x + t + a_2)^{a_3} \right) - \frac{1}{a_3}, -\ln a_1(x + a_2)^{a_3} \right\}
- b_1b_2^{2b_2} \left[ \frac{(x + t)^{-2b_2 + 1}}{-2b_2 + 1} - \frac{x^{-2b_2 + 1}}{-2b_2 + 1} \right]
- \int_{0}^{t} \frac{c_1c_2^{(x+s)^{\gamma_3}}}{1 + c_1c_2^{(x+s)^{\gamma_3}}} \, ds \right\} = e^0 = 1 , \]

\[\lim_{t \to \infty} S_{x}^{HP4}(t) = \lim_{t \to \infty} \exp \left\{ -\frac{1}{a_3} \left( \frac{-1}{\ln a_1} \right)^{\frac{1}{\gamma_3}} \left( \frac{1}{a_3}, -\ln a_1(x + t + a_2)^{a_3} \right) - \frac{1}{a_3}, -\ln a_1(x + a_2)^{a_3} \right\}
- b_1b_2^{2b_2} \left[ \frac{(x + t)^{-2b_2 + 1}}{-2b_2 + 1} - \frac{x^{-2b_2 + 1}}{-2b_2 + 1} \right]
- \int_{0}^{t} \frac{c_1c_2^{(x+s)^{\gamma_3}}}{1 + c_1c_2^{(x+s)^{\gamma_3}}} \, ds \right\} = 0 . \]

We examine the first part of the above limit.
\[
\lim_{t \to \infty} \exp \left\{ - \frac{1}{a_3} \left( - \frac{1}{\ln a_1} \right)^{\frac{1}{a_3}} \left( \gamma \left( \frac{1}{a_3} , - \ln a_1 (x+t+a_2)^{a_3} \right) - \gamma \left( \frac{1}{a_3} , - \ln a_1 (x+a_2)^{a_3} \right) \right) \right\} = \\
\exp \left( \frac{1}{a_3} \left( - \frac{1}{\ln a_1} \right)^{\frac{1}{a_3}} \gamma \left( \frac{1}{a_3} , - \ln a_1 (x+a_2)^{a_3} \right) \right) \lim_{t \to \infty} \exp \left( - \frac{1}{a_3} \left( - \frac{1}{\ln a_1} \right)^{\frac{1}{a_3}} \times \gamma \left( \frac{1}{a_3} , - \ln a_1 (x+t+a_2)^{a_3} \right) \right) \\
\times \gamma \left( \frac{1}{a_3} , - \ln a_1 (x+t+a_2)^{a_3} \right) = \exp \left( \frac{1}{a_3} \left( - \frac{1}{\ln a_1} \right)^{\frac{1}{a_3}} \gamma \left( \frac{1}{a_3} , - \ln a_1 (x+a_2)^{a_3} \right) \right) \times 0 = 0 .
\]

where \( \gamma(a,t) = (a - 1) \gamma(a - 1, t) - t^{a-1} e^{-t} \) and \( \lim_{t \to \infty} t^{a} e^{-t} = 0 \) [14, 23].

For the second part, using the Limit Chain Rule, — reader can refer to the book [14] — we get

\[
g(t) = (x+t)^{-2b_2+1} , \\
f(u) = e^u ,
\]

Taking the \( \lim_{t \to \infty} g(t) = \infty \) we get:

\[
\lim_{t \to \infty} -b_1 b_3^{2b_2} \left[ \frac{(x+t)^{-2b_2+1}}{-2b_2 + 1} - \frac{x^{-2b_2+1}}{-2b_2 + 1} \right] = \\
- \frac{b_1 b_3^{2b_2}}{-2b_2 + 1} \lim_{t \to \infty} \left( (x+t)^{-2b_2+1} - x^{-2b_2+1} \right) = \\
- \frac{b_1 b_3^{2b_2}}{-2b_2 + 1} \left( \infty - x^{-2b_2+1} \right) = -\infty
\]

and

\[
\lim_{u \to -\infty} e^u = 0
\]

Combining all the parts, both \( S_{x}^{HP4}(t) \) and \( F_{x}^{HP4}(t) \) have proper boundary behavior.

\[\square\]
2.3 Power-exponential Function Based Model

Even though it has been recently proposed, the power-exponential function based model is promising and can adequately describe the mortality pattern. The model is proposed by Lundengård et al [16] and is based on the use of power-exponential functions as defined in [23]. Using these functions as building blocks, we can build a model with \( n \) “humps”. In this project, we use a model for a single hump.

Let \( A = \{(a_1, a_2, a_3)\} \) and we get the expected hump when \( a_i > 0 \) for \( i = 1, 2, 3 \).

**Definition 4** (Power-exponential Function Based Model — PE). The mathematical expression of the model is defined as

\[
\mu(x) = \frac{c_1}{xe^{-c_2x}} + a_1 \left(xe^{-a_2x}\right)^{a_3}.
\]  

(2.15)

Each parameter can easily be interpreted in terms of qualitative properties of the curve.

For a further discussion of the particular model as well as the derivation for its survival function for multiple humps along with its properties, the reader can refer to [16]. Following, after fitting the data to the model, the five parameter estimates for the year 2010 are given in the table below and a graphic representation of the resulting mortality rate curve of the Power-exponential function based model is given in Fig. 2.7.

As seen in Eq. (2.15), the model contains various parameters that modify the shape of the mortality curve for the Power-exponential model. Following, we provide different figures that show the effect on the mortality rate curve when a unique parameter is varied, while the other parameters retain the values given in Table 2.3.

From Fig. 2.8 we can observe the parameters that have a significant effect on the initial curve. First of all, the parameter \( c_2 \) approximately gives the slope of the logarithm of the mortality rate curve for the older ages. In addition, the parameter \( a_1 \) shows the different heights that the “accident hump” can take while the parameter \( a_2 \) shows the time when the “accident hump” will occur. For example, as the value of \( a_1 \) and \( a_2 \) increases the height of the hump decreases and the sooner the hump occurs respectively. Finally, the parameter \( a_3 \) determines the steepness of the “accident hump”.

**Theorem 5** (PE – Survival Function). *Given the Theorem 2 we get that the survival function*

Table 2.3: Parameter estimates for the logarithm of the mortality rate fitted to mortality in the Swedish population 2010.

<table>
<thead>
<tr>
<th>( c_1 )</th>
<th>( c_2 )</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( a_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0002</td>
<td>0.1230</td>
<td>7.5071</td>
<td>0.0399</td>
<td>22.7592</td>
</tr>
</tbody>
</table>

24
Figure 2.7: Example of the mortality rate curve given by the Power-exponential Model in the year 2010.
Figure 2.8: The effect on the logarithm of the mortality rate curve when one parameter is varied, while the others retain the values given in Table 2.3.
of Power-exponential function based model is:

\[
S_{PE}^x(t) = \exp \left\{ c_1 \left( \text{Ei}(c_2 x) - \text{Ei}(c_2 (x + t)) \right) \\
+ \frac{a_1}{(a_2 a_3)^{a_3 + 1}} \left( \gamma(a_3 + 1, a_2 a_3 x) - \gamma(a_3 + 1, a_2 a_3 (x + t)) \right) \right\},
\]  \hspace{1cm} (2.16)

where \( \text{Ei}(x) = - \int_{-\infty}^{x} \frac{e^{-s}}{s} \, ds \) is the exponential integral and \( \gamma(a, t) = \int_{0}^{t} x^{a-1} e^{-x} \, dx \) is the lower incomplete Gamma function \([1]\).

The survival function satisfies the two conditions:

1. \( S_{PE}^x(0) = 1 \),
2. \( \lim_{t \to \infty} S_{PE}^x(t) = 0 \).

**Proof.** Once again, for computational convenience we will calculate

\[
\ln S_{PE}^x(t) = \int_{0}^{t} \frac{c_1}{(x + s)e^{-c_2(x+s)}} + a_1 \left( (x + s)e^{-a_2(x+s)} \right)^{a_3} \, ds.
\]

We calculate each part separately. We set \( u = c_2(x+s) \) so we get \( du = c_2 \, ds \).

i) \[
\int_{0}^{t} \frac{c_1}{(x + s)e^{-c_2(x+s)}} \, ds = \int_{c_2 x}^{c_2(x+t)} \frac{c_1}{ue^{-u}} \, du = c_1 \text{Ei}(u) \bigg|_{c_2 x}^{c_2(x+t)} = c_1 \left( \text{Ei}(c_2(x+t)) - \text{Ei}(c_2 x) \right).
\]  \hspace{1cm} (2.17)

ii) Here we set \( u = a_2 a_3(x+s) \leftrightarrow du = a_2 a_3 \, ds \). So we get:
\[ \int_0^t a_1 (x+s) e^{-a_2(x+s)} \] \[ a_3 \] \[ ds = \int_{a_2a_3x}^{a_2a_3(x+t)} \left( \frac{u}{a_2a_3} \right)^{a_3} e^{-a_1/a_2a_3} du \]

\[ = \frac{a_1}{(a_2a_3)^{a_3+1}} \int_{a_2a_3x}^{a_2a_3(x+t)} u^{a_3} e^{-u} \, du \]

\[ = \frac{a_1}{(a_2a_3)^{a_3+1}} \left[ \gamma(a_3+1, \, u) \right]_{a_2a_3x}^{a_2a_3(x+t)} \]

\[ = \frac{a_1}{(a_2a_3)^{a_3+1}} \left( \gamma(a_3+1, \, a_2a_3(x+t)) - \gamma(a_3+1, \, a_2a_3x) \right) . \]

(2.18)

Combining Eq. (2.17) and Eq. (2.18) we get Eq. (2.16).

\[ \square \]

In the next chapter, we explain the selected method for forecasting mortality rates known as the Lee–Carter model.
Chapter 3

Selected Methods for Forecasting

3.1 The Lee–Carter Model

Due to the rapid aging of the population, a more important issue have been brought to the fore; forecasting mortality rates. The methods for forecasting are distributed into three main approaches: 1) explanation — forecasting based on structural or epidemiological models — 2) expectation — forecasting based on expert opinion — and 3) extrapolation — forecasting based on patterns and trends. Many recent researchers highlight that the most promising approach for long-term mortality forecasting is those methods based on extrapolation. For that reason, Lee and Carter create a principal components approach that could be used for the extrapolation of trends and age patterns in making long-run mortality forecasting [2, 3, 17].

**Definition 5** (Lee–Carter Model). This well-known method is based on a combination of statistical time series methods and we denote:

- \( \mu_{x,t} \), the mortality rate at age \( x \) in year \( t \),
- \( a_x \), the average log-mortality at age \( x \),
- \( b_x \), the rate of change at age \( x \) in response to the level of mortality over time,
- \( k_t \), the level of mortality in year \( t \), and
- \( \varepsilon_{x,t} \), the residual.

Thus, the two-factor Lee–Carter model is given by:

\[
\ln(\mu_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t},
\]  

(3.1)
Following, $a_x$, $b_x$ and $k_t$ have to be normalized. However, the below properties need to be satisfied:

$$a_x = \frac{\sum \ln(m_{x,t})}{T},$$  \hspace{1cm} (3.2)

where $T$ is the overall time,

$$\sum b_x = 1,$$  \hspace{1cm} (3.3)

$$\sum k_t = 0.$$  \hspace{1cm} (3.4)

The $k_t$ is an adjusted estimation and using statistical time series methods we can forecast $k_t$ index by:

$$k_t = k_{t-1} + d + e_t,$$  \hspace{1cm} (3.5)

where $d$ is the drift and $e_t$ is the error [6, 17, 24].

The reader can also refer to the original paper from Lee and Carter for further explanation in the model and its properties [3, 13, 17]. The process that Lee–Carter model follows four steps. First of all, using the Singular Value Factorization (SVF), — known also as Singular Value Decomposition (SVD) — the Lee–Carter model can find a unique least squares solution using Eq. (3.1), and thus forecast age-specific mortality rates [6]. Specifically, the $a_x$, $b_x$ and $k_t$ can be estimated using the Least Squares method. The Least Square method is used to find the best fitted version of the model and we explain it extensively in section 3.2 [6].

In what follows, we use standard univariate methods — Autoregressive Integrated Moving Average model (ARIMA) — Eq. (3.5) to forecast the mortality index $k_t$ and, along with the vectors $a_x$ and $b_x$, we can adjust it by fitting it to each year’s life expectancy [2, 17]. In fact, due to its linearity, $k_t$ vector accurately captures the trends of mortality rates implying that the ratio of the rates are fairly constant across time in most populations at different ages [3]. However, our models will be fitted to data for measured death rates from many different countries using numerical techniques for curve-fitting with the non-linear least squares method (see Section 3.2).

Finally, the Lee–Carter model is commonly used for long-term forecasting due to its satisfying results as well as its advantages. Many researchers highlight that the Lee–Carter is now the dominant method for forecasting mortality — mainly, in developed countries — because of its simplicity and accuracy. In fact, the parameters of the model are very few comparing to other methods and thus they can be easily interpreted and through the use of a random walk (Eq. (3.5)) the model can forecast mortality and smoothly displays it in an individual’s life span [3, 24]. Moreover, another strength of the Lee–Carter model is that it adopts the main characteristic of the methods based on extrapolation capturing better the log mortality trends especially in the case that they are linear [24].

In other words, the model’s minimal subjectivity enhances the mortality forecast despite the non-linearity in the trends and produces a very good fit to mortality trends due to its constraints on the way the rates will behave in the future. However, one disadvantage of the model is the assumption that the ratio of the rates are fairly constant across time in most populations at
different ages [3]. Nevertheless, modifications of the Lee–Carter have been introduced and they provide more flexibility in forecasting change performing as well as the initial approach. In the following chapter, we analyze the mortality data and forecast mortality rates using the initial Lee–Carter approach.

3.2 Regression and the Least Square Method

In this section, we discuss about the least square method which is used in applications of numerical techniques such as the curve fitting of experimental data. In fact, the Least Square method is used to find the “best” fitted version of the model and for that reason we define what the method is below. We consider:

- **A** will be a matrix that has equal number of rows, \(n\), and columns, \(m\), — denoted as \(A_{n \times m}\) — and
- **x, y** will be the vectors representing the points on a straight line.

In that case, we have a square matrix \(A\) that has a unique solution.

**Definition 6** (Least Square Method). We observe a system of equations in the matrix form \(Ax = y\) such that \(A_{n \times m}\) with \(n > m\) (an over-determined system). If the residual vector is defined as:

\[
e = Ax - y
\]

then minimizing its Euclidean length:

\[
e^2 = \sum_{i=1}^{n} |e|^2 = \sqrt{e^\top e} = \sqrt{(Ax - y)^\top (Ax - y)}.
\]

we obtain the best fit in the least-squares sense of the model used to generate the starting system of equations. The Least Square method can be used both in linear and non-linear regression problems.

3.2.1 Linear Regression

First of all, we will discuss a scenario where we have a system of linear equations.

**Definition 7** (Linear Regression). We consider a function in the form:

\[
f(x) = \sum_{j=1}^{m} r_{j} g_{j}(x_{ij})
\]
and
\[
\mathbf{r} = [r_1 \ r_2 \ \ldots \ r_m]^\top
\]
as a set of parameters, where \(x_{ij}\) for all \(i = 1, 2, \ldots, n\) and \(j = 1, 2, \ldots, m\) are variables and \(g_j\) are functions.

Suppose we have a set of values:

- \(\mathbf{x} = [x_1 \ x_2 \ \ldots \ x_m]^\top\) and
- \(\mathbf{y} = [y_1 \ y_2 \ \ldots \ y_n]^\top\).

The corresponding system of equations will be

\[
\mathbf{X} \mathbf{r} = \mathbf{y},
\]
where

\[
\mathbf{X} = \begin{bmatrix}
g_1(x_{11}) & g_2(x_{12}) & \ldots & g_m(x_{1m}) \\
g_1(x_{21}) & g_2(x_{22}) & \ldots & g_m(x_{2m}) \\
\vdots & \vdots & \ddots & \vdots \\
g_1(x_{n1}) & g_2(x_{n2}) & \ldots & g_m(x_{nm})
\end{bmatrix},
\quad \mathbf{r} = \begin{bmatrix}
r_1 \\
r_2 \\
\vdots \\
r_m
\end{bmatrix},
\quad \mathbf{y} = \begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}.
\]

Thus, residual vector will be:

\[
\mathbf{e} = \mathbf{X} \mathbf{r} - \mathbf{y} = \begin{bmatrix}
e_1 \\ e_2 \\ \vdots \\ e_n
\end{bmatrix} \tag{3.8}
\]

The reader can refer also to the following references for a more detailed explanation [6, 19]. For illustration, our equation looks like \(y(x) = a + bx\). The corresponding system of equations will be:

\[
\mathbf{X} \mathbf{r} = \mathbf{y},
\]
where

\[
\mathbf{X} = \begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
\vdots & \vdots \\
1 & x_n
\end{bmatrix},
\quad \mathbf{r} = \begin{bmatrix}
a \\ b
\end{bmatrix},
\quad \mathbf{y} = \begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}.
\]

Following, the vector \(\mathbf{e}\) will be:

\[
\mathbf{e} = \mathbf{X} \mathbf{r} - \mathbf{y} = \begin{bmatrix}
a + bx_1 - y_1 \\ a + bx_2 - y_2 \\ \vdots \\ a + bx_n - y_n
\end{bmatrix}.
\]
3.2.2 Non–Linear Regression

However, there is a possibility that we might have a non-linear regression problem. This type of regression is used when the models we want to fit is not linear with respect to their parameters [6, 19].

**Definition 8** (Non–Linear Regression). We consider a function with \( m \) variables. We define:

- \( \mathbf{x} = [x_1 \ x_2 \ \ldots \ x_m]^\top \) as set of parameters,
- \( \mathbf{r} = [r_1 \ r_2 \ \ldots \ r_k]^\top \) and
- function \( f(\mathbf{x}; \mathbf{r}) \).

Suppose we have a set of values \( x_i, y_i \), where \( i = 1, 2, \ldots, n \). The corresponding system of equations will be

\[
\mathbf{X} \ \mathbf{r} = \mathbf{y}.
\]

Thus, residual vector will be:

\[
\mathbf{e} = \mathbf{X} \ \mathbf{r} - \mathbf{y} = \begin{bmatrix}
  f(x_1; r) - y_1 \\
  f(x_2; r) - y_2 \\
  \vdots \\
  f(x_n; r) - y_n
\end{bmatrix}
\]

(3.9)

For example, our equation looks like \( y(x) = ae^{bx} \). Even though we have non-linear model, we can turn it into linear problem by \( \Rightarrow \ln y(x) = \ln a + bx \). The corresponding system of equations will be:

\[
\mathbf{X} \ \mathbf{r} = \mathbf{y},
\]

where

\[
\mathbf{X} = \begin{bmatrix}
  1 & x_1 \\
  1 & x_2 \\
  \vdots & \vdots \\
  1 & x_n
\end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix}
  \ln a \\
  b
\end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix}
  \ln y_1 \\
  \ln y_2 \\
  \vdots \\
  \ln y_n
\end{bmatrix}.
\]

It is important to point out that it is usually more difficult to find \( \mathbf{r} \) comparing to a linear regression case and it often requires a numerical method to minimise the norm of the residual vector [6]. This can be done using the Eq. (3.7). In the next section, the models referred in Chapter 2 will be fitted to data from Human Mortality Database for measured death rates from many different countries using numerical techniques for curve-fitting with the non-linear least squares method.
Chapter 4

Model Fitting

In this chapter, we analyse the mortality data. We gathered data from the Human Mortality Database regarding the mortality rate data from different countries and we fit the three models from Chapter 2 to the mortality rate data using standard least-square curve-fitting techniques on $\ln(\mu)$ assuming constant variance, see also Chapter 3 [12]. Therefore, in this paper we select to examine and fit the mortality rate data for USA, Sweden and Greece in different years and comparing the results we make valid conclusions in terms of the life expectancy in each country as well as how well the models fit the mortality rates, see Section 4.1. Furthermore, in Section 4.2, following the process discussed in Chapter 3 and using the Lee–Carter model we forecast the mortality rates of each country for different periods and we compare the performance of all models with respect to the actual data. Additionally, part of those results will be presented during the 5th International Conference on Stochastic Modeling Techniques and Data Analysis, SMTDA 2018 that will take place in Chania, Crete, Greece on the 12th – 15th of June 2018 [15]. A code in MATLAB is implemented to generate the graphs provided by Karl Lundengård and it can be obtained upon request.

4.1 Analysis of Mortality Data

We get the data regarding the mortality rates of USA, Sweden and Greece and fit the three models — Modified Perks, Heligman–Pollard, and Power-exponential — which gives us results shown in the following figures. Specifically, we examine the mortality rates for male population because a vast majority of studies have shown that the “accident hump” for female population is decreasing and as the years pass by it disappears [3, 11, 24].

From Figs. 4.1, 4.2 and 4.3, we can observe how well each model performs in respect with the expected mortality rates of the male population. It is important to note that the Modified Perks model in all countries does not perform well comparing to the other models, as expected,
Figure 4.1: Mortality rates fitted by the three models for male population of USA. Figure 4.1a is the population for the year 1970, Figure 4.1b is population for the year 2000 and Figure 4.1c is population for the year 2010.
Figure 4.2: Mortality rates fitted by the three models for male population of Sweden. Figure 4.2a is population for the year 1970, Figure 4.2b is population for the year 2000 and Figure 4.2c is population for the year 2010.
Figure 4.3: Mortality rates fitted by the three models for male population of Greece. Fig. 4.3a is population for the year 1981, Fig. 4.3b is population for the year 2000 and Fig. 4.3c is population for the year 2010.
because it cannot show the hump that occurs during the adulthood. For example, in Figure 4.3 when the magnitude of the hump is big or affects a greater range of ages then the Modified Perks cannot fit the data well even in ages 40+ where the Gompertz law is applied. Among the other two models, HP4 and Power-exponential, even though both of them performs well to the expected mortality rates, the first one has a “better” fit to that well-known hump. It fact, Power-exponential model underestimates the mortality rates in infant stage by making a more smooth curve. That might be caused due to less number of parameters in the equation. However, the two models have been proven to be an excellent model on matching the shape of mortality rates in the younger ages and after the age of forty.

In the case of USA, we can conclude that there is a decline in accident hump in the recent years, see also Figs. 4.1a and 4.1c. That would be the effect of fast developing countries in the hump due to improvement in medical industry and lifestyle. Comparing the first two figures — Fig. 4.2a and Fig. 4.2b — we can highlight that there is an improvement in the life expectancy of the male population during the first years of their life. However, in the year 1970 we could see that the “accident hump” is smaller in magnitude and spread than the earlier years and that is because the old times there was less traffic and thus less accidents, see also Fig. 4.2. Finally, comparing the last two figures that represent the mortality rates in Greece — Fig. 4.3a and Fig. 4.3b — we get more or less the same results but the difference among the decrease of the hump and the improvement in life expectancy is less obvious than that of Sweden. Consequently, we can conclude that among the three aforementioned mortality models the HP4 model performs better and has a “better” fit to the data that represent the three key periods of mortality in an individual’s life span: infant mortality, accident hump, and senescent mortality. In the next section, we forecast mortality rates using the Lee–Carter model and compare the results of the three models with the actual data.

### 4.2 Forecasting Mortality Rates

In this section, we fit the three aforementioned models to the data for male population from the Human Mortality Database using standard least-square curve-fitting techniques. We generate new data using the fitted models and then we predict mortality rates for selected years applying the Lee–Carter method explained in Chapter 3. In the following figures, we forecast mortality rates of USA, Sweden and Greece for ten years making solid conclusions on the performance of each model on the data. In addition, we examine the perfect ratio on data availability so we can get the most accurate forecasting.

From the following figures we can see that among the three models Heligman–Pollard model and Power-exponential model are the ones that perform better and accurately display the mortality rate curve of each country. In terms of data availability and comparing the results during the selected periods, we used mortality data of twenty-five years for both USA and Sweden to forecast mortality rates for the next fifteen years, while for Greece we used mortality data of twenty years to forecast mortality rates for the next ten years. It is important to notice that
Figure 4.4: Forecasting mortality rates for male population of USA in the years 2000–2010. The figure displays the forecasting for the last year.

Figure 4.5: Forecasting mortality rates for male population of Sweden in the years 2000–2010. The figure displays the forecasting for the last year.
Figure 4.6: Forecasting mortality rates for male population of Greece in the years 2001–2011. The figure displays the forecasting for the last year.

how well a model performs is based on the data availability; that can be seen if we compare both the Figures 4.4 and 4.5 with the Figure 4.6. In fact, all models perfectly fit the curve after the age 40 but not in case of Greece. However, among the Heligman–Pollard model and Power-exponential model the first one due to the number of parameters has a “better” fit in all countries in all stages of an individual’s lifespan especially during the adulthood where the “accident” hump occurs.
Chapter 5

Conclusion

5.1 Project summary

In this project, we focus on the risk involved in insurance companies and more specifically the financial impact of the unexpected mortality in that sector. For that reason, over the last decades, the modeling of a law of mortality has been a consistent interest from a vast majority of researchers [11, 16]. The goal of this thesis is to systematically evaluate a few different models to determine their relative strengths and weaknesses with respect to forecasting the mortality rate.

In the first part of the thesis, after a brief review of existing models for mortality rates, we discussed about the selective mortality models — namely the Modified Perks model, the Heligman–Pollard (HP4) model and the Power-exponential model — and we analysed how each model’s mortality curve looks like. Following, we discussed the different methods for forecasting mortality and focused on the Lee-Carter model which we used for forecasting and measuring the quality of the forecast.

In the second part of the thesis, data for measured death rates for male population of USA, Sweden and Greece have been fitted by three selected models using numerical techniques for curve-fitting with the non-linear least squares method. Then, after generating new data using the three fitted models, we forecasted mortality rates of the three countries for ten years using the Lee–Carter method.

Therefore, we can conclude that the Heligman–Pollard (HP4) model and the Power-exponential model accurately display mortality rate curve for male population. However, regarding the forecast of mortality rates the Heligman-Pollard model has a “better” fit in all stages of an individual’s lifespan especially during the adulthood where the “accident hump” occurs.
5.2 Future Work

In this thesis, we compared three different mortality rates model in terms of how well they display mortality curve. Even though both the Heligman–Pollard (HP4) model and the Power-exponential model perform better than the Modified Perks model, the former outperforms the latter one. That is due to the larger number of parameters that can capture the mortality curve better regardless the time consumption needed. However, the Power-exponential model adequately performs with a smooth demonstration of the curve during the infant mortality stage despite the fact that it is recently developed. Thus, for further research, the Power-exponential model can be promising if there will be an extensive study on the parameters so it can model the “accident hump” properly as well as exploring other methods of forecasting in combination with the Power-exponential model.
Chapter 6

Summary of reflection of objectives in the thesis

A summary of the objectives accomplished in this thesis will be presented in this chapter.

6.1 Objective 1: Knowledge and understanding

This thesis demonstrates the knowledge and understanding applied mathematics as well as the concepts related to actuarial mathematics, applied matrix analysis and numerical methods. In Chapter 1 and 3, a vast majority of sources related to the subject has been used to extensively describe the different mortality models along with the various methods to forecast mortality. In Chapter 2, the understanding of the theory based on the force of mortality is required. Moreover, computer skills are required for writing the thesis. Even though it was the first time, the student searched many relevant sources to properly write the thesis using the \LaTeX.

6.2 Objective 2: Methodological knowledge

After the theoretical description, the student thoroughly presented the subject and demonstrated a methodological knowledge through the use of figures making the topic more understandable for the reader. This is successfully achieved using the MATLAB. One of the major challenges was the understanding of the code due to the fact that the student was not familiar with the particular program; nevertheless, previous experience of similar programs along with the support of supervisor Karl Lundengård help to overcome this challenge.
6.3 Objective 3: Critically and Systematically Integrate Knowledge

Information from different sources has been used in the thesis. First of all, many sources have been suggested from supervisor Milica Rančić. During the process, the student found a vast majority of journal articles as well as books to extensively discuss and elaborate on the particular concept.

6.4 Objective 4: Independently and Creatively Identify and Carry out Advanced Tasks

The first stage of the thesis was to do a research on the topic. After familiarizing with it, the student decided different chapters and stages that the thesis should include: i) literature review, ii) theoretical results and iii) empirical results. Finally, the student made valid conclusions based on the results obtained. Discussions and improvements of the thesis with the supervisors Milica Rančić and Karl Lundengård on relevant parts were very valuable for the quality of the end result.

6.5 Objective 5: Present and Discuss Conclusions and Knowledge

The structure of the thesis makes it easy for the readers unfamiliar with the subject to understand the topic. Using many sources the student extensively explained all subjects related to the topic. Figures and tables are provided for the reader’s easy understanding regarding the comparisons and several aspects of the results that are discussed in the thesis. Those who want a deeper understanding can refer to the last part of the thesis that includes the list of sources used as well as the algorithm used to obtain results. An oral presentation of the thesis will take place in May 2018 where the audience will have the chance to attend and ask questions.

6.6 Objective 6: Scientific, Social and Ethical Aspects

All the sources used for the theoretical description are properly cited as well as all the data marked with citations. The algorithm is provided by supervisor Karl Lundengård. Finally, all those who supported in the completion of the thesis have been thanked in the Acknowledgments.
Bibliography


