Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Hok-Kwan Kan

A Bayesian Mortality Forecasting Framework for Population and Portfolio Mortality

MSc Thesis 2012-035

UNIVERSITEIT VAN AMSTERDAM

A Bayesian Mortality Forecasting Framework for Population and Portfolio Mortality

A thesis submitted to the
University of Amsterdam
Faculty of Economics and Business
in partial fulfillment of the requirements

for the degree

Master of Science
in
Actuarial Science and Mathematical Finance

by

Hok Kwan Kan (6002978)

The Hague, The Netherlands
11th January 2012

Primary supervisor: dr. Katrien Antonio
Secondary supervisor: Prof.dr.ir. Michel Vellekoop

Abstract

The life expectancy in industrialized countries has increased remarkably in recent decades. While most individuals consider this as a positive development, increased life expectancy has huge implications for governments, pension funds and insurance companies. Therefore actuaries and demographers make use of mortality models to forecast future mortality. Furthermore, the upcoming Solvency II regulations, which are due by 2013, require European insurance companies to quantify the uncertainty around the forecast of future mortality, i.e. mortality risk. Uncertainty quantification requires the need of models with stochastic features.

Since its introduction in 1992 the Lee-Carter model has become the leading stochastic mortality model in the actuarial and demographic literature. Subsequently many extensions to the standard Lee-Carter model have been proposed. This thesis studies in particular the extension of the Lee-Carter model to the Poisson-gamma setting. In comparison with the commonly used Poisson approach, the Poisson-gamma approach can explicitly capture unexplained variability in the data by introducing dispersion parameters. This thesis compares the two approaches using Dutch population mortality data.

Insurance companies and pension funds do not face mortality risk related to a population, but mortality risk related to their portfolio. However, most stochastic mortality models cannot be fit reliably for portfolio mortality, because the amount of historical portfolio data is limited in terms of the size of the dataset as well as the number of years of portfolio data. In this thesis, we propose an extension to the Poisson-gamma Lee-Carter model for portfolio mortality estimation. The proposed extension applies Bayesian inference techniques based on the conjugacy of the Poisson and gamma distributions. Additionally, the proposed extension allows using the full forecasting ability of the Lee-Carter model. The mathematics behind the extension is closely related to credibility theory more commonly applied in the field of non-life insurance. We illustrate the results using the Dutch life and pension portfolio of the insurance company AEGON.

Table of Contents

Chapter 1

Introduction

1.1 Background

In the twentieth century the life expectancy in the Netherlands increased remarkably. In 1900 we expected a newborn male infant to live 47 years, while if the same male infant was to be born in 2008 we would expect him to live 78.3 years. Most developed countries experienced similar improvement in life expectancy and there are no signs yet that further improvement of life expectancy will come to a halt soon. The latter development is being recognized in recent publications. Statistics Netherlands (van Duin & Garssen, 2010) estimates the life expectancy of a male infant born in 2060 to rise to 84.5 years and the Dutch Actuarial Association (*Prognosetafel 2010-2060*, 2010) estimates an increase to 85.9 years.

While most individuals consider increased life expectancy as a positive development, it has huge implications for governments, such as the public retirement planning. Pension funds, insurance companies and other providers of mortality-linked products are also affected, because lower mortality immediately affects their pricing and reserving. Actuaries use mortality tables which contain current mortality statistics and often also forecasts of future mortality. Thus the implications caused by increased life expectancy are taken into account as long as the realized mortality does not differ significantly from the forecasted mortality. This stresses the importance of a 'good' mortality table and mortality model, but production of such a table and model does not come easily.

The upcoming Solvency II regulations, which are due by 2013, require European insurance companies to hold a certain amount of capital on their balance sheets. This amount of capital should be sufficient to cover future cash flows as well as to absorb risks to protect policyholders and beneficiaries. Therefore in addition to the forecast of future mortality, the Solvency II regulations require the quantification of the uncertainty around the forecast, which is referred to as mortality risk. More specifically, the term longevity risk is used when focus is on the risk that people tend to live longer than expected. The quantification of mortality risk requires stochastic mortality models.

Note that insurance companies and pension funds do not face mortality risk related to a population, but mortality risk related to their portfolio. However, most stochastic mortality models cannot be fit reliably for portfolio mortality, because the amount of historical portfolio data is limited in terms of the size of the dataset as well as the number of years of portfolio data.

1.2 Objective and motivation

Since its introduction in 1992 the Lee-Carter model (Lee & Carter, 1992) has become the leading stochastic mortality model in the actuarial and demographic literature. Subsequently many extensions to the standard Lee-Carter model have been proposed. Particularly the extensions of the Lee-Carter model to a Poisson-gamma setting considered by Delwarde *et al.* (2007) and Li *et al.* (2009) caught our interest. The key difference between the two approaches centers on the dispersion parameter. Delwarde *et al.* (2007) propose a general dispersion parameter, while Li *et al.* (2009) introduce age-specific dispersion parameters. Our main focus lies on the latter model and therefore, unless otherwise noted, the use of the term "Poisson-gamma" in this text refers to the Poisson-gamma approach with age-specific dispersion parameters. In this thesis we explore the applicability of the Poisson-gamma Lee-Carter model to the Dutch population mortality. The standard Poisson Lee-Carter model is hereby used as a benchmark.

Next, we set ourselves the goal to contribute to the research of stochastic forecasting models for portfolio mortality. Olivieri & Pitacco (2009) proposed an approach to determine a portfolio experience factor with Bayesian inference techniques using the conjugacy of the Poisson and gamma distributions. We recognize that such an approach can be applied to the Poisson-gamma Lee-Carter model in a natural way. Based on this approach we propose a Bayesian extension for the Poisson-gamma Lee-Carter model which can be used to estimate portfolio mortality and more importantly allows using the full forecasting ability of the Lee-Carter model for portfolio mortality. Moreover, we discuss the connection between this approach and the credibility framework, known from non-life insurance. We explore the applicability of this extension using the Dutch life and pension portfolio of AEGON.

1.3 Outline of the thesis

The remainder of the thesis consists of two parts. The first part compares the performance of the Poisson-gamma Lee-Carter model to the standard Poisson Lee-Carter model in the context of Dutch population mortality. Chapter 2 describes the standard Lee-Carter model with the traditional linear regression approach as well as the Poisson regression approach. Since our main focus lies on the extension of the Lee-Carter model to a Poisson-gamma setting with age-specific dispersion parameters, we introduce in detail this approach first in chapter 3, while chapter 4 briefly describes the Poisson-gamma approach with general dispersion parameter. In chapter 5 we apply all three models to the Dutch population mortality, present their results and compare their performance.

The second part starts with chapter 6 where we present a Bayesian extension to the Poisson-gamma Lee-Carter model for portfolio mortality based on the conjugate property of the Poisson-gamma distribution. In chapter 7 we apply the Bayesian extension to the Dutch life and pension portfolio of AEGON and present the results. Finally in chapter 8 we present the conclusions and contributions of the preceding chapters.

Chapter 2

The Lee-Carter model

In 1992 Ronald Lee and Lawrence Carter published their seminal work on forecasting models for human mortality (Lee & Carter, 1992). Since then the Lee-Carter model has been widely adopted for diverse actuarial and demographic applications. For instance, the United States Census Bureau used the model as a benchmark in their population forecast (Hollmann *et al.*, 2000). Many extensions to the standard Lee-Carter model have been proposed, for example to account for age-time interactions (Booth *et al.*, 2002), age-specific cohort effects (Renshaw & Haberman, 2006) and age-specific enhancements (Renshaw & Haberman, 2003a).

This chapter is organized as follows. Section 2.1 first introduces to the reader the notation and specification of the Lee-Carter model. Section 2.2 presents the traditional linear regression approach of the Lee-Carter model and the singular value decomposition to obtain the ordinary least-squares solution. Section 2.3 describes the Lee-Carter model approached as a Poisson regression model along with the numerical methods to obtain the maximum-likelihood solution. Section 2.4 discusses the forecasting of mortality with the Lee-Carter model using standard time series methods. Section 2.5 discusses simulation methods to obtain confidence intervals for the mortality forecast.

2.1 The model specification

The central death rate $m_{x,t}$ denotes the average death rate experienced within a group of population aged x in year t. We use the following definition of the central death rate:

$$m_{x,t} = \frac{d_{x,t}}{e_{x,t}} \tag{2.1}$$

where $d_{x,t}$ and $e_{x,t}$ stand for the number of deaths and respectively the average number of people living in year t aged x. The latter $e_{x,t}$ is also referred to as the exposure. For an observed population $e_{x,t}$ is often estimated by taking the population at mid-year. We refer to the natural logarithm of the central death rate as the log-mortality $\ln(m_{x,t})$.

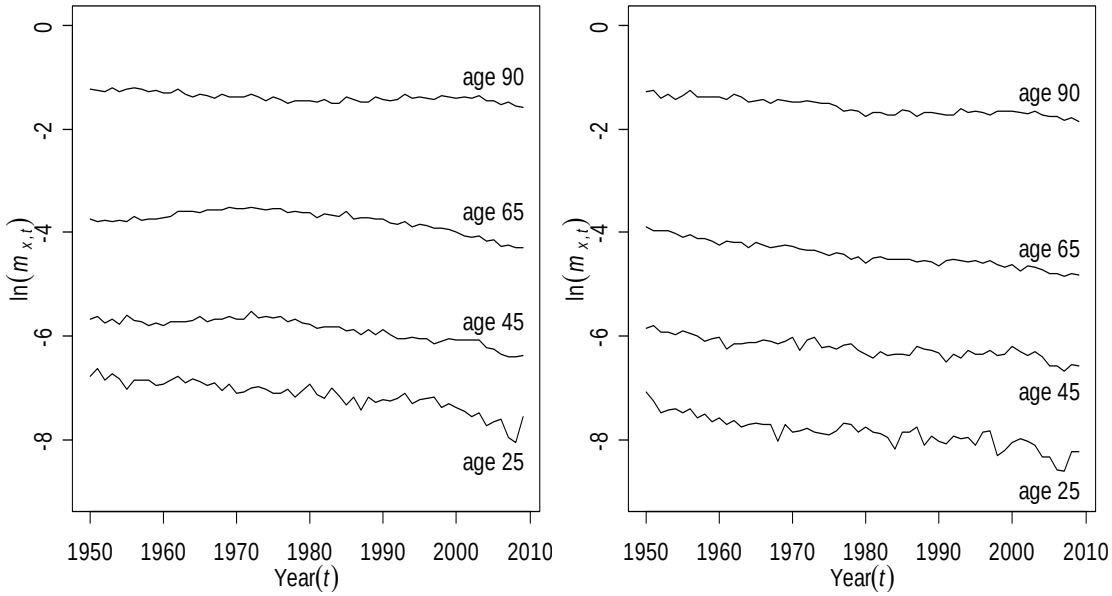

Figure 2.1: Historical log-mortality development of Dutch males (left panel) and females (right panel)

Figure 2.1 shows the age-specific development over time of the log-mortality for Dutch males and females. We observe that for each age the log-mortality starts at a different level and that for all ages the log-mortality tends to decline linearly over time, though not necessarily with equal slope. We will see that the rationale behind the Lee-Carter model is closely related to the observed development of the Dutch log-mortality.

Lee & Carter (1992) proposed modeling the central death rate at age x and time t as follows

$$\ln(m_{x,t}) = \alpha_x + \beta_x \kappa_t + \epsilon_{x,t} \tag{2.2}$$

The $\{\alpha_x\}$ coefficients describe the overall level of mortality corresponding with a specific age. The $\{\kappa_t\}$ coefficients represent the time trend reflecting the general level of mortality in time, also referred to as the mortality index. The $\{\beta_x\}$ coefficients reflect the age-specific sensitivity to changes in the mortality index and the product of β_x and κ_t reflects the age-specific development of the mortality level in time. This relation becomes apparent when taking the derivative with respect to t.

$$\frac{d(\alpha_x + \beta_x \kappa_t)}{dt} = \beta_x \frac{d\kappa_t}{dt} \tag{2.3}$$

The error terms $\varepsilon_{x,t}$ capture the remaining variability which cannot be explained by the model.

The model is undetermined under the following transformations for any constant c

$$\{\alpha_x, \beta_x, \kappa_t\} \;\rightarrow\; \{\alpha_x, \beta_x / c\,, c \cdot \kappa_t\} \tag{2.4}$$

$$\{\alpha_x, \beta_x, \kappa_t\} \;\rightarrow\; \{\alpha_x - c \cdot \beta_x, \beta_x, \kappa_t + c\} \tag{2.5}$$

Thus different parameterizations exist which form equivalent formulations. To make the model identifiable constraints need to be imposed on the parameters. In their original paper Lee and Carter proposed the identification constraints

$$\sum_x \beta_x = 1 \tag{2.6}$$

$$\sum_t \kappa_t = 0 \tag{2.7}$$

The second constraint implies that α_x equals the average of $\ln(m(x, t))$ over time, which can be seen as follows

$$\sum_{t=t_1}^{t_n} \ln(m_{x,t}) = \sum_{t=t_1}^{t_n} (\alpha_x + \beta_x \kappa_t + \epsilon_{x,t}) = (t_n - t_1 + 1)\alpha_x + \beta_x \sum_{t=t_1}^{t_n} \kappa_t + \sum_{t=t_1}^{t_n} \epsilon_{x,t} \tag{2.8}$$

with t_1 and t_n respectively denoting the first and last observed year. Applying the identification constraint (2.7) and assuming that the average of the error terms $\varepsilon_{x,t}$ equals zero and vanishes, lead to the following result.

$$\alpha_x = \frac{\sum_{t=t_1}^{t_n} \ln(m_{x,t})}{t_n - t_1 + 1} \tag{2.9}$$

The Lee-Carter model is simple and parsimonious, which makes the model intuitive and easily understandable. Additionally, the Lee-Carter model reduces the mortality development over time for all ages to one single time trend, the mortality index. Hence, the difficulty of forecasting future mortality levels is now reduced to forecasting the mortality index. However, mortality as a function of one single time trend also implies perfect correlation between changes in mortality at all ages

$$Cor(\alpha_{x_1} + \beta_{x_1}\kappa_t, \alpha_{x_2} + \beta_{x_2}\kappa_t) = 1 \; \forall x_1, x_2, t \tag{2.10}$$

which does not seem biologically reasonable.

2.2 Estimating Lee-Carter using OLS

Traditionally the Lee-Carter model has been approached as a linear regression model where the error terms are independent and identically distributed with mean 0 and variance σ_ε^2. In the left hand side of equation (2.2), the observed log-mortality $\ln(m(x, t))$ acts as the response variable. The right hand side of that equation however does not contain any observed regressors. The estimation of the parameters takes place using ordinary least-squares (OLS), hence minimizing

$$\sum_{x,t} (\ln(m_{x,t}) - \alpha_x + \beta_x \kappa_t)^2 \tag{2.11}$$

The lack of regressors prevents us from using the familiar regression methods. Therefore Lee & Carter (1992) resort to the singular value decomposition (SVD) to obtain the OLS solution.

OLS estimation with SVD

First the $\{\alpha_x\}$ coefficients are estimated as in equation (2.9) which implicitly follows from applying the identification constraint (2.8).

$$\hat{\alpha}_x = \frac{\sum_{t=t_1}^{t_n} \ln(m_{x,t})}{t_n - t_1 + 1} \tag{2.12}$$

Then the estimation problem reduces to finding the estimates for $\{\beta_x\}$ and $\{\kappa_t\}$ which minimize (2.11). By applying SVD one obtains the decomposition

$$\ln(m_{x,t}) - \hat{\alpha}_x = \sum_{i=1}^{r} \rho_i U_{x,i} V_{t,i} \tag{2.13}$$

where $r = \text{rank}[\ln(m_{x,t}) - \hat{\alpha}_x]$ and ρ_i for i = 1,…, r are the ordered (increasingly) singular values with $U_{x,i}$ and $V_{t,i}$ as the corresponding left and right singular vectors. Using only the first eigenvalue results in the following approximation

$$\hat{\beta}_x \hat{\kappa}_t = \rho_1 U_{x,1} V_{t,1} \tag{2.14}$$

while imposing both the identification constraints (2.6) and (2.7).

Note that the estimation takes place on the log-mortality, i.e. $\ln(m_{x,t})$, instead of the central death rate. Furthermore, the expected number of deaths, obtained by applying the estimated central death rate to the observed population, does not necessarily equal the observed number of deaths. To establish this desirable property the coefficients $\{\hat{\kappa}_t\}$ are re-estimated while keeping $\{\hat{\alpha}_t\}$ and $\{\hat{\beta}_t\}$ fixed such that

$$\sum_x d_{x,t} = \sum_x e_{x,t} \hat{m}_{x,t} \tag{2.15}$$

with $\hat{m}_{x,t} = e^{\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t}$

Jump-off bias correction

The estimated death rate in the last observed year does not (necessarily) equal the observed death rate in that year. Hence, the following condition does not necessarily hold

$$\ln(m_{x,t_n}) = \ln(\hat{m}_{x,t_n}) \tag{2.16}$$

where t_n denotes the last observed year. When forecasting this will cause a discontinuity between the last observed death rate and the first forecasted death rate. This discontinuity is often referred to as the jump-off bias. Lee & Carter (1992) accepted this discontinuity stating that the jump-off bias affects only death rates which are absolutely very low and have little impact on the forecasted life expectancy. However, Bell (1997) as well as Lee & Miller (2001) concluded that a correction for the jump-off bias improves the forecast of life expectancy, especially in the early years of the forecast. The jump-off bias can easily be corrected for by modifying the estimation for $\{\hat{\alpha}_t\}$ and the identification constraint (2.7) to

$$\hat{\alpha}_x = \ln(m_{x,t_n}) \tag{2.17}$$

$$\kappa_{t_n} = 0 \tag{2.18}$$

Approaching the Lee-Carter model as a linear regression model gives the observed death rates at younger ages the same weight as older ages, even though the older ages contain less observations. To solve this problem Wilmoth (1993) proposed fitting the Lee-Carter model using weighted least-squares (WLS) instead of OLS. The linear

regression approach also implies homoskedastic error terms, but this does not agree with our observation of the death rates. At older ages the observed death rates show more variability than at younger ages, which can be explained by the fact that the number of lives as well as the number of deaths at older ages is much smaller. The next section discusses a modification of the Lee-Carter model allowing heteroskedastic error terms.

2.3 Estimating Lee-Carter using Poisson regression and maximum likelihood

Assuming that the death counts can be approximated by a Poisson distribution (often contributed to Brillinger (1986)) Brouhns *et al.* (2002a) approached the Lee-Carter model as a Poisson regression model.

$$D_{x,t} \sim \text{Poisson}\left(e_{x,t} m_{x,t}\right) \tag{2.19}$$

with $m_{x,t} = e^{\alpha_x + \beta_x \kappa_t}$. The main advantage of this modification is the introduction of heteroskedastic error terms. Renshaw & Haberman (2003b) also proposed using the Poisson error structure, with time as a known covariate. However, Brouhns *et al.* (2002a) modeled time as a factor.

Instead of resorting to SVD the estimation of the parameters takes place using the maximum likelihood (ML) method, i.e. maximizing the log-likelihood of model (2.19)

$$\begin{aligned} l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}; \boldsymbol{d}) &= \log \prod_{x,t} f\left(d_{x,t}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}\right) = \log \prod_{x,t} e^{-e_{x,t} m_{x,t}} \frac{\left(e_{x,t} m_{x,t}\right)^{d_{x,t}}}{d_{x,t}!} \\ &= \sum_{x,t} \left[-e_{x,t} m_{x,t} + d_{x,t} \cdot \ln\left(e_{x,t} m_{x,t}\right) - \ln\left(\Gamma\left(d_{x,t} + 1\right)\right) \right] \end{aligned} \tag{2.20}$$

Due to the presence of the bilinear term $\beta_x \kappa_t$ the proposed model cannot be implemented as an ordinary GLM and therefore most statistical software packages cannot fit the model using standard routines.

One-dimensional maximization using Newton's method

To maximize the (log-)likelihood we resort to Newton's method, a numerical method for finding the roots of a real-valued function. Searching for the maximum or minimum of a real-valued function *f(x)* implies looking for the roots of the derivative

$$f'(x) = 0 \tag{2.21}$$

Applying a Taylor series to (2.21) implies

$$f'(x + h) \approx f'(x) + f''(x) \cdot h \tag{2.22}$$

which is a linear function in h that approximates f near a given x. The root of the linear approximation in h can easily be determined to be

$$h = -\frac{f'(x)}{f''(x)} \tag{2.23}$$

assuming that $f''(x) \neq 0$. Hence, the root of the first order Taylor approximation to f at x is

$$x + h = x - \frac{f'(x)}{f''(x)} \tag{2.24}$$

where h is known as the Newton step. This motivates the iterative updating scheme shown in Algorithm 2.1. For more detailed information about finding roots with numerical methods we refer the interested reader to Heath (2002).

```
x ← initial guess
do
    x ← x − f′(x)/f″(x)
until x converges
```

Algorithm 2.1: Newton's method for one-dimensional non-linear functions

Multidimensional maximization with Newton's method

Algorithm 2.1 can be extended for multi-dimensional non-linear functions $f(\boldsymbol{x})$. However, this requires the calculation of the Hessian matrix consisting of second order partial derivatives of $f(\boldsymbol{x})$

$$\left\{H_f(\boldsymbol{x})\right\}_{ij} = \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} \tag{2.25}$$

Goodman (1979) proposed a simpler method for estimating log-linear models with bilinear terms. The method consists of iteratively applying the one-dimensional Newton's method to each dimension of f separately in a round-robin fashion until convergence is reached. Algorithm 2.2 describes the method in pseudo code.

```
x ← initial guess
do
    for i ← 1, 2, …, k do
        apply Newton's method to f(…,x_i, …)
    end
until x converges
```

Algorithm 2.2: Applying the one-dimensional Newton method to a k-dimensional non-linear function

The log-likelihood function (2.20) has three sets of parameters $\{\alpha_x\}$, $\{\beta_x\}$ and $\{\kappa_t\}$. Let us denote by $\hat{\alpha}_x^{(i)}$, $\hat{\beta}_x^{(i)}$ and $\hat{\kappa}_x^{(i)}$ the estimates at iteration i with initial values $\hat{\alpha}_x^{(0)} = 0$, $\hat{\beta}_x^{(0)} = 1$ and $\hat{\kappa}_x^{(0)} = 0$. In a next step the parameter sets are updated using Goodman's algorithm. Note that after each update of $\{\beta_x\}$ and $\{\kappa_t\}$ the parameters need to be rescaled to satisfy identification constraints (2.6) and (2.7). The complete updating scheme is as follows with $\hat{m}_{x,t}^{(i)} = e^{\hat{\alpha}_x^{(i)} + \hat{\beta}_x^{(i)}\hat{\kappa}_x^{(i)}}$

(a) Update $\{\alpha_x\}$:
$$\begin{cases} \hat{\alpha}_x^{(i+1)} = \hat{\alpha}_x^{(i)} - \frac{\sum_t \left(d_{x,t} - e_{x,t}\hat{m}_{x,t}^{(i)}\right)}{-\sum_t e_{x,t}\hat{m}_{x,t}^{(i)}} \ \forall x \\ \hat{\beta}_x^{(i+1)} = \hat{\beta}_x^{(i)} \ \forall x \\ \hat{\kappa}_t^{(i+1)} = \hat{\kappa}_t^{(i)} \ \forall t \end{cases}$$

(b) Update $\{\beta_x\}$: $\begin{cases} \hat{\alpha}_x^{(i+2)} = \hat{\alpha}_x^{(i+1)} \ \forall x \\ \tilde{\beta}_x^{(i+2)} = \hat{\beta}_x^{(i+1)} - \frac{\sum_t \hat{\kappa}_t^{(i+1)}\left(d_{x,t} - e_{x,t}\hat{m}_{x,t}^{(i+1)}\right)}{-\sum_t e_{x,t}\hat{m}_{x,t}^{(i+1)}\left(\hat{\kappa}_t^{(i+1)}\right)^2} \ \forall x \\ \tilde{\kappa}_t^{(i+2)} = \hat{\kappa}_t^{(i+1)} \ \forall t \end{cases}$

(c) Apply identification constraint (2.6): $\begin{cases} \hat{\beta}_x^{(i+2)} = \tilde{\beta}_x^{(i+2)} / \sum_x \tilde{\beta}_x^{(i+2)} \ \forall x \\ \hat{\kappa}_t^{(i+2)} = \tilde{\kappa}_t^{(i+2)} \sum_x \tilde{\beta}_x^{(i+2)} \ \forall t \end{cases}$

(d) Update $\{\kappa_t\}$: $\begin{cases} \tilde{\alpha}_x^{(i+3)} = \hat{\alpha}_x^{(i+2)} \ \forall x \\ \hat{\beta}_x^{(i+3)} = \hat{\beta}_x^{(i+2)} \ \forall x \\ \tilde{\kappa}_t^{(i+3)} = \hat{\kappa}_t^{(i+2)} - \frac{\sum_x \hat{\beta}_x^{(i+2)}\left(d_{x,t} - e_{x,t}\hat{m}_{x,t}^{(i+2)}\right)}{-\sum_x e_{x,t}\hat{m}_{x,t}^{(i+2)}\left(\hat{\beta}_x^{(i+2)}\right)^2} \ \forall t \end{cases}$

(e) Apply identification constraint (2.7): $\begin{cases} \hat{\alpha}_x^{(i+3)} = \tilde{\alpha}_x^{(i+3)} + \frac{\sum_{t=t_1}^{t_n} \tilde{\kappa}_t^{(i+3)}}{t_n - t_1 + 1} \hat{\beta}_x^{(i+3)} \ \forall x \\ \hat{\kappa}_t^{(i+3)} = \tilde{\kappa}_t^{(i+3)} - \frac{\sum_{t=t_1}^{t_n} \tilde{\kappa}_t^{(i+3)}}{t_n - t_1 + 1} \ \forall t \end{cases}$

Jump-off bias correction

The model fitted according to the estimation procedure described above also suffers from the jump-off bias described in section 2.2. The jump-off bias can be corrected for by the modification shown in equation (2.17) and (2.18). To incorporate the correction for the jump-off bias steps (a) and (e) of the updating scheme need to be replaced by

(a) Update $\{\alpha_x\}$ **once**: $\begin{cases} \hat{\alpha}_x^{(i+1)} = \ln m_{x,t_n} \ \forall x \\ \hat{\beta}_x^{(i+1)} = \hat{\beta}_x^{(i)} \ \forall x \\ \hat{\kappa}_t^{(i+1)} = \hat{\kappa}_t^{(i)} \ \forall t \end{cases}$

(e) Apply identification constraint (2.18): $\begin{cases} \hat{\alpha}_x^{(i+3)} = \tilde{\alpha}_x^{(i+3)} + \tilde{\kappa}_{t_n}^{(i+3)} \hat{\beta}_x^{(i+3)} \ \forall x \\ \hat{\kappa}_t^{(i+3)} = \tilde{\kappa}_t^{(i+3)} - \tilde{\kappa}_{t_n}^{(i+3)} \ \forall t \end{cases}$

2.4 Mortality forecasting

The Lee-Carter model reduces the mortality development over time for all ages to one single time trend, the mortality index. Hence, the problem of forecasting future mortality is now reduced to forecasting the mortality index κ_t. Lee & Carter (1992) proposed using an appropriate ARIMA(p, d, q) time series model, which takes the general form

$$\begin{cases} \nabla^d \kappa_t = v_t + c \\ \varphi(L) v_t = \theta(L) \epsilon_t \end{cases} \tag{2.26}$$

where

- L is the lag operator, i.e. $L\kappa_t = \kappa_{t-1}$ and $L^s \kappa_t = L^{s-1}\kappa_{t-1}$.
- ∇ is the difference operator, i.e. $\nabla\kappa_t = \kappa_t - \kappa_{t-1}$ and $\nabla^d \kappa_t = \nabla^{d-1}\kappa_t - \nabla^{d-1}\kappa_{t-1}$. $\nabla^d \kappa_t$ is also referred to as a time series with integration of order d.
- $\varphi(L)$ is the autoregressive polynomial of degree p with coefficients $\varphi_1, \ldots, \varphi_p$, i.e. $\varphi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \varphi_3 L^3 - \cdots - \varphi_p L^p$.

- $\theta(L)$ is the moving average polynomial of degree q with coefficients $\theta_1, \ldots, \theta_q$, i.e. $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \theta_3 L^3 + \cdots + \theta_q L^q$.
- ϵ_t is a white noise process, independent identical distributed as $\epsilon_t \sim N(0, \sigma_\epsilon^2)$.
- c is the intercept which for $d = 0$ captures the mean of the time series and for $d > 0$ captures the drift in the time series.

An appropriate ARIMA model is found by carrying out the standard Box-Jenkins methodology consisting of the following four stages.

Model identification

First the stationarity of the variable needs to be determined, e.g. by applying an augmented Dickey–Fuller (ADF) unit-root test. A generally applied method to achieve stationarity is differencing the data once or multiple times. In a next step the orders of the autoregressive and moving average polynomial have to be determined. Each time series model has a unique theoretical autocorrelation function (ACF) and a partial autocorrelation function (PACF). By (visually) comparing the theoretical ACF and PACF with the sample ACF and PACF it is possible to identify suitable candidates for the ARIMA model.

Model estimation

In the next stage the parameters of the potential ARIMA models need to be estimated. Measures such as the "Akaike Information Criterion" (AIC) and "Bayesian Information Criterion" (BIC) can help in selecting the final ARIMA model. Both criteria are based on the residual variance. Since the residual variance decreases by adding more parameters, both criteria also incorporate a penalty to discourage overfitting. The AIC and BIC are defined as

$$AIC = -2\ln(L) + 2k \tag{2.27}$$

$$BIC = -2\ln(L) + k\ln(n) \tag{2.28}$$

where L denotes the maximized value of the likelihood function, k the number of parameters and n the number of observations used to estimate the model. Hence, the model with the lowest information criterion is the preferred model.

Diagnostic checking

Once the ARIMA model has been fitted to the data, the absence of residual autocorrelation needs to be checked. An ARIMA model is considered correctly specified if the residuals form a white noise process. A well-known test for the absence of autocorrelation is the Ljung-Box Q-test, which under the null hypothesis of uncorrelated residuals defines the following Q-statistic

$$Q = n(n+2)\sum_{k=1}^{K}\frac{\hat{\rho}_k^2}{n-k} \overset{H_0}{\sim} \chi^2(K-p-q) \tag{2.29}$$

where $\hat{\rho}_k$ denotes the residual autocorrelation coefficient of lag length k, and K is the number of lags being tested.

Forecasting

After having found an appropriate ARIMA model the variable, in our case the mortality index κ_t, can be forecasted. Let us denote the s-period ahead forecast of the mortality index as $\hat{\kappa}_{t_n+s}$. Then in case of the Poisson Lee-Carter model, the expected value of future death count is given by

$$E\left[D_{x,t_n+s}\right] = e_{x,t_n+s}\hat{m}_{x,t_n+s} \tag{2.30}$$

where e_{x,t_n+s} is the future exposure and $\hat{m}_{x,t_n+s}$ is the forecast of future death rate with

$$\hat{m}_{x,t_n+s} = e^{\hat{\alpha}_x+\hat{\beta}_x\hat{\kappa}_{t_n+s}} \tag{2.31}$$

Using $\hat{m}_{x,t_n+s}$ we can calculate other quantities of interest, such as life expectancies, life annuity premiums, etc.

2.5 Simulation and confidence intervals

The previous section discussed the methodology to obtain forecasts for the expected values (so-called 'best estimates') of quantities of interest. In forecasting, information on the uncertainty of the forecasted quantities plays an even more important role. The uncertainty directly relates to the risk an insurance company or pension fund faces. Confidence intervals act as a useful tool to describe the uncertainty around a best estimate. When forecasting with the Lee-Carter model uncertainty arises from three sources:

- Forecast uncertainty, due to the stochastic error in the forecast of κ_t;
- Parameter uncertainty, due to the sampling error in the estimated parameters of the model;
- Model uncertainty, related to the sensitivity of the outcome due to model selection.

In our thesis we do not focus on model uncertainty. For a discussion on this subject we refer the interested reader to Cairns (2000).

In the original Lee-Carter method confidence intervals are based only on the forecast error. Concerning prediction of life expectancy Lee & Carter (1992) concluded that the forecast uncertainty dominates over the parameter uncertainty for long-run forecasts as well as short-run forecasts. In this case they found it acceptable to ignore the sampling error when deriving confidence intervals. For death rates forecasts, they concluded that confidence intervals based on forecast error alone provide only reasonable results for forecast horizons greater than 10 to 25 years. For actuarial applications in general we would like to measure the uncertainty of all sources. The forecast errors of an ARIMA model can be analytically derived, but determining confidence intervals for the combined forecast and sampling errors still remains analytically intractable. Another difficulty comes from the quantities of interest (e.g. future mortality levels, life expectancies and life annuity premiums) being complicated non-linear functions of the Lee-Carter model parameters ($\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\kappa}$) and the ARIMA model parameters (c, $\boldsymbol{\varphi}$, $\boldsymbol{\theta}$, σ_ε). Therefore we need to resort to simulation techniques.

In the literature many simulation techniques have been suggested to capture both the forecast uncertainty as well as the parameter uncertainty. Brouhns *et al.* (2002b) proposed a parametric bootstrap approach using the following property. When using maximum

likelihood to estimate the parameters of the Poisson Lee-Carter model (2.19), then $(\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\kappa}})$ is asymptotically multivariate normally (MVN) distributed with mean $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa})$ and covariance matrix given by the inverse of the Fisher information matrix. Later Brouhns *et al.* (2005) suggested a semi-parametric bootstrap approach which consists of generating bootstrap samples of the death counts $d_{x,t}$ by drawing from the estimated Poisson distribution given by (2.19). The distribution of $(\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\kappa}})$ is obtained by re-estimating the Poisson Lee-Carter model using the bootstrap samples. Koissi *et al.* (2006) applied a residual (non-parametric) bootstrap approach. In this approach a bootstrap sample of $d_{x,t}$ is obtained by first sampling with replacement from the deviance residuals, and finally the corresponding bootstrap sample of $d_{x,t}$ is obtained by applying the inverse formula for the deviance residual. For a comparison of the three simulation techniques we refer the interested reader to Renshaw & Haberman (2008).

Simulation algorithm

In our thesis we apply the semi-parametric bootstrap approach of Brouhns *et al.* (2005) which consists of the following steps:

1. Generate *N* bootstrap samples of the death count $d_{x,t}^{(n)}$ for $n = 1, \dots, N$, where $d_{x,t}^{(n)}$ are realizations from the Poisson distribution given by (2.19).
2. For the *n*th bootstrap sample:
   (a) Re-estimate the Poisson Lee-Carter model obtaining the parameter sets $\{\alpha_x^{(n)}\}$, $\{\beta_x^{(n)}\}$, $\{\kappa_t^{(n)}\}$;
   (b) Using $\{\kappa_t^{(n)}\}$ re-estimate the parameters of the originally estimated ARIMA model where the order of integration *d*, the degree *p* of the autoregressive polynomial and the degree *q* of the moving average polynomial do not change;
   (c) Generate a projection of future mortality index $\kappa_t^{(n)}$ using the ARIMA model obtained in step (2b);
   (d) Calculate the forecasted future mortality rate;
   (e) Calculate the quantity of interest using the forecasted future mortality rate.
3. Derive the confidence interval for the quantity of interest using the empirical distribution obtained in step (2).

This approach has several appealing features. It takes into account both the parameter uncertainty (in step 1) as well as the forecast uncertainty (in step 2). The approach can easily be modified for other distributions than the Poisson distribution, and extended for more complex models.

Chapter 3

The Lee-Carter model in a Poisson-gamma setting with age-specific dispersion parameters

Section 2.3 introduced model (2.19) which approaches the Lee-Carter model as a Poisson regression model. The assumption of the Poisson model has the drawback of imposing a mean-variance equality restriction on $D_{x,t}$.

$$E\left[D_{x,t}\right] = Var\left[D_{x,t}\right] \tag{3.1}$$

If the empirical data show evidence that the variance is larger or smaller than expected according to the model, the data are called overdispersed respectively underdispersed. To account for overdispersion Renshaw & Haberman (2006) resort to the quasi-Poisson distribution. Delwarde *et al.* (2007) follow an alternative approach to account for overdispersion based on the negative binomial distribution, which can be viewed as the generalization of the Poisson model to a Poisson-gamma model. Li *et al.* (2009) also propose a model based on the negative binomial distribution. The key difference between the two approaches centers on the dispersion parameter. Delwarde *et al.* (2007) propose a general dispersion parameter, while Li *et al.* (2009) introduce an age-specific dispersion parameter. Chapter 4 describes the former approach, while this chapter discusses the latter approach which we mainly follow in our thesis.

Section 3.1 provides an overview of the mathematical properties of the Poisson-gamma distribution. Section 3.2 discusses the extension of the Lee-Carter model to the Poisson-gamma setting and section 3.3 presents the numerical methods to obtain the corresponding maximum-likelihood. Section 3.4 discusses the forecasting of mortality with the Poisson-gamma Lee-Carter model. Section 3.5 presents a simulation method to obtain the confidence intervals for the mortality forecast.

3.1 Introducing the Poisson-gamma setting

The Poisson distribution is commonly used to represent count data. This distribution is especially appealing in cases when a limited amount of data is available, because the Poisson distribution needs only one parameter to be estimated. The drawback of the simple and parsimonious Poisson distribution arises from the implicitly imposed mean-

variance equality restriction (3.1). To overcome this restriction and account for overdispersion the Poisson distribution can be generalized as follows

$$\begin{cases} N|\Lambda \sim Poisson(\Lambda) \\ \Lambda \sim Gamma(\alpha, \beta) \end{cases} \tag{3.2}$$

Such a generalization, where the parameter of the Poisson distribution is a random variable itself, is referred to as a mixed Poisson distribution or Poisson-gamma distribution. The random variable Λ, which acts as the parameter of the Poisson distribution, is called the structural variable. The unconditional mean and variance of this distribution are

$$E[N] = E\big[E[N|\Lambda]\big] = E[\Lambda] = \frac{\alpha}{\beta} \tag{3.3}$$

$$Var(N) = E[Var(N|\Lambda)] + Var(E[N|\Lambda]) = E[\Lambda] + Var(\Lambda) \geq E[\Lambda] \tag{3.4}$$

Equation (3.4) shows that the Poisson-gamma distribution can account for overdispersion and that it is not restricted by the mean-variance equality. The unconditional distribution of N can be derived with the moment-generating function in the following way.

$$\begin{aligned} m_N(t) &= E[e^{tN}] = E\big[E[e^{tN}|\Lambda]\big] = E\big[e^{\Lambda(e^t-1)}\big] = m_\Lambda(e^t - 1) \\ &= \left(\frac{\beta}{\beta-(e^t-1)}\right)^\alpha = \left(\frac{p}{1-(1-p)e^t}\right)^\alpha \text{ with } p = \frac{\beta}{\beta+1} \end{aligned} \tag{3.5}$$

The latter form can be recognized as the moment-generating function of the negative binomial distribution $NegBin\left(\alpha, \frac{\beta}{\beta+1}\right)$. For applications of the mixed Poisson distribution, commonly used within non-life insurance, we refer to Kaas *et al.* (2008).

3.2 The model specification

The Poisson Lee-Carter model assumes homogeneity within an age-period cell, i.e. the death rate $m_{x,t}$ applies to all individuals aged x at time t. However, this assumption seems to be invalid, because most likely an age-period cell will contain individuals having different backgrounds, for instance in ethnicity, education, occupation, marital status etc.

To incorporate more heterogeneity into the Lee-Carter model each age-period cell is divided into N_x equal clusters each having exposure $e_{x,t}^{(i)} = \frac{e_{x,t}}{N_x}$ and death count $D_{x,t}^{(i)}$ for $i = 1, \dots, N_x$. Assuming independency between $D_{x,t}^{(i)}$ and $D_{x,t}^{(j)}$ for $i \neq j$, Li *et al.* (2009) model the death count $D_{x,t}^{(i)}$ as a Poisson-gamma model.

$$\begin{cases} D_{x,t}^{(i)}|Z_x^{(i)} \sim Poisson\left(e_{x,t}^{(i)} m_{x,t} Z_x^{(i)}\right) \\ \quad Z_x^{(i)} \sim Gamma(a_x, b_x) \\ \quad \text{with } m_{x,t} = e^{\alpha_x + \beta_x \kappa_t} \end{cases} \tag{3.6}$$

Intuitively, the gamma distributed random variable $Z_x^{(i)}$ acts as the experience factor, thus assigning death rate $m_{x,t}Z_x^{(i)}$ to cluster i. Further, we impose the assumption

$$E\left[Z_x^{(i)}\right] = 1 \tag{3.7}$$

which implies that on average each cluster has the same death rate as the age-period cell they belong to. Using assumption (3.7) we reparameterize the gamma distribution to

$$Z_x^{(i)} \sim Gamma(\varphi_x^{-1}, \varphi_x^{-1}) \tag{3.8}$$

The distribution for the total number of deaths $D_{x,t}$ can be derived by summing the death counts over the clusters. We use the property that the sum of independent Poisson distributions also leads to a Poisson distribution where the parameter equals the sum of the component parameters.

$$\begin{aligned} D_{x,t} | Z_x^{(1)}, \dots, Z_x^{(N_x)} &= \sum_{i=1}^{N_x} D_{x,t}^{(i)} | Z_x^{(i)} \sim Poisson\left(\sum_{i=1}^{N_x} e_{x,t}^{(i)} m_{x,t} Z_x^{(i)} \right) \\ &= Poisson\left(e_{x,t} m_{x,t} \frac{\sum_{i=1}^{N_x} Z_x^{(i)}}{N_x} \right) = Poisson\left(e_{x,t} m_{x,t} \bar{Z}_x \right) \end{aligned} \tag{3.9}$$

Next, the distribution of $\bar{Z}_x$ equals the average of $Z_x^{(i)}$ for $i = 1, \dots, N_x$. Hence,

$$\bar{Z}_x = \frac{\sum_{i=1}^{N_x} Z_x^{(i)}}{N_x} \sim Gamma(N_x \varphi_x^{-1}, N_x \varphi_x^{-1}) = Gamma(\bar{\varphi}_x^{-1}, \bar{\varphi}_x^{-1}) \tag{3.10}$$

Summarizing the results above, the Lee-Carter model can be approached as a Poisson-gamma model in the following manner

$$\begin{cases} D_{x,t} | \bar{Z}_x \sim \text{Poisson}\left(e_{x,t} m_{x,t} \bar{Z}_x \right) \\ \bar{Z}_x \sim Gamma(\bar{\varphi}_x^{-1}, \bar{\varphi}_x^{-1}) \end{cases} \tag{3.11}$$

If $\bar{Z}_x \sim Gamma(\bar{\varphi}_x^{-1}, \bar{\varphi}_x^{-1})$, then $e_{x,t} m_{x,t} \bar{Z}_x \sim Gamma\left(\bar{\varphi}_x^{-1}, \left(e_{x,t} m_{x,t} \bar{\varphi}_x \right)^{-1} \right)$. Using the results derived in section 3.1 we conclude for the unconditional distribution of $D_{x,t}$

$$D_{x,t} \sim NegBin\left(\bar{\varphi}_x^{-1}, \frac{1}{1 + e_{x,t} m_{x,t} \bar{\varphi}_x} \right) \tag{3.12}$$

which has unconditional mean and variance

$$E\left[D_{x,t} \right] = e_{x,t} m_{x,t} \tag{3.13}$$

$$Var\left(D_{x,t} \right) = e_{x,t} m_{x,t} + \bar{\varphi}_x \left(e_{x,t} m_{x,t} \right)^2 \tag{3.14}$$

Note that if $\bar{\varphi}_x \to 0$ the Poisson-gamma model reverts back to the Poisson model. This can be seen by taking the limit of the probability mass function of (3.12) resulting in the probability mass function of the Poisson distribution.

$$\begin{aligned}
&\lim_{\bar{\varphi}_x\to 0}\binom{\bar{\varphi}_x^{-1}+d_{x,t}-1}{d_{x,t}}\left(\frac{1}{1+e_{x,t}m_{x,t}\bar{\varphi}_x}\right)^{\bar{\varphi}_x^{-1}}\left(1-\frac{1}{1+e_{x,t}m_{x,t}\bar{\varphi}_x}\right)^{d_{x,t}}\\
&=\lim_{\bar{\varphi}_x\to 0}\frac{\Gamma\left(\bar{\varphi}_x^{-1}+d_{x,t}\right)}{\Gamma(\bar{\varphi}_x^{-1})d_{x,t}!}\left(\frac{1}{1+e_{x,t}m_{x,t}\bar{\varphi}_x}\right)^{\bar{\varphi}_x^{-1}}\left(\frac{e_{x,t}m_{x,t}}{\bar{\varphi}_x^{-1}+e_{x,t}m_{x,t}}\right)^{d_{x,t}}\\
&=\frac{\left(e_{x,t}m_{x,t}\right)^{d_{x,t}}}{d_{x,t}!}\lim_{\bar{\varphi}_x\to 0}\frac{\Gamma\left(\bar{\varphi}_x^{-1}+d_{x,t}\right)}{\Gamma(\bar{\varphi}_x^{-1})\left(\bar{\varphi}_x^{-1}+e_{x,t}m_{x,t}\right)^{d_{x,t}}}\lim_{\bar{\varphi}_x\to 0}\left(\frac{1}{1+e_{x,t}m_{x,t}\bar{\varphi}_x}\right)^{\bar{\varphi}_x^{-1}}\\
&=\frac{\left(e_{x,t}m_{x,t}\right)^{d_{x,t}}}{d_{x,t}!}\cdot 1\cdot\frac{1}{e^{e_{x,t}m_{x,t}}}=e^{-e_{x,t}m_{x,t}}\frac{\left(e_{x,t}m_{x,t}\right)^{d_{x,t}}}{d_{x,t}!}
\end{aligned}\tag{3.15}$$

The generalization to the Poisson-gamma Lee-Carter model has the appealing feature that the mean equals the mean of the Poisson Lee-Carter model for equal $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\kappa}$, as shown by equation (3.13). More importantly, equation (3.14) states that the Poisson-gamma approach explicitly allows for overdispersion using an age-specific dispersion parameter $\bar{\varphi}_x$ and therefore is able to capture the heterogeneity within an age-period cell.

3.3 Estimation by maximum likelihood

To estimate the Lee-Carter model as given by (3.11), Li *et al.* (2009) resort to the maximum likelihood method. This requires the maximization of the corresponding log-likelihood function given by

$$\begin{aligned}
l(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa},\bar{\boldsymbol{\varphi}};\boldsymbol{d}) &= \log\prod_{x,t}f\left(d_{x,t};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa},\bar{\boldsymbol{\varphi}}\right)\\
&=\log\prod_{x,t}\binom{\bar{\varphi}_x^{-1}+d_{x,t}-1}{d_{x,t}}\left(\frac{1}{1+e_{x,t}m_{x,t}\bar{\varphi}_x}\right)^{\bar{\varphi}_x^{-1}}\left(1-\frac{1}{1+e_{x,t}m_{x,t}\bar{\varphi}_x}\right)^{d_{x,t}}\\
&=\sum_{x,t}\left\{\sum_{i=0}^{d_{x,t}-1}\log(\bar{\varphi}_x^{-1}+i)-\log\left(d_{x,t}!\right)+d_{x,t}\log\left(e_{x,t}m_{x,t}\bar{\varphi}_x\right)\right.\\
&\qquad\left.-\left(d_{x,t}+\bar{\varphi}_x^{-1}\right)\log\left(1+e_{x,t}m_{x,t}\bar{\varphi}_x\right)\right\}
\end{aligned}\tag{3.16}$$

Multidimensional maximization with Newton's method

Maximizing the log-likelihood function takes place using the algorithm of Goodman (1979) as described in section 2.3. Denoting $\hat{\alpha}_x^{(i)}$, $\hat{\beta}_x^{(i)}$, $\hat{\kappa}_x^{(i)}$ and $\hat{\bar{\varphi}}_x^{(i)}$ as the estimates at iteration i, we first define the following functions:

- $\hat{d}_{x,t}^{(i)}=e_{x,t}\widehat{m}_{x,t}=e_{x,t}e^{\hat{\alpha}_x^{(i)}+\hat{\beta}_x^{(i)}\hat{\kappa}_x^{(i)}}$;
- $f_{x,t}^{(i)}=\sum_{i=0}^{d_{x,t}-1}\left(\frac{i}{1+\hat{\bar{\varphi}}_x^{(i)}\cdot i}-\frac{1}{\hat{\bar{\varphi}}_x^{(i)}}\right)$;
- $g_{x,t}^{(i)}=\hat{d}_{x,t}^{(i)}\left(d_{x,t}+\frac{1}{\hat{\bar{\varphi}}_x^{(i)}}\right)\left(1+\hat{\bar{\varphi}}_x^{(i)}\hat{d}_{x,t}^{(i)}\right)$;

- $h_{x,t}^{(i)} = \sum_{i=0}^{d_{x,t}-1}\left(\frac{-i^2}{\left(1+\hat{\bar{\varphi}}_x^{(i)}\cdot i\right)^2} - \frac{1}{\left(\hat{\bar{\varphi}}_x^{(i)}\right)^2}\right)$;
- $r_{x,t}^{(i)} = \frac{\log\left(1+\hat{\bar{\varphi}}_x^{(i)}\hat{d}_{x,t}^{(i)}\right)}{\left(\hat{\bar{\varphi}}_x^{(i)}\right)^2}$.

Next, the iterative scheme consists of the following updating steps:

(a) Update $\{\alpha_x\}$: $\begin{cases}\hat{\alpha}_x^{(i+1)} = \hat{\alpha}_x^{(i)} - \frac{\sum_t\left(d_{x,t}-\hat{\bar{\varphi}}_x^{(i)}g_{x,t}^{(i)}\right)}{-\sum_t\left(-\hat{\bar{\varphi}}_x^{(i)}g_{x,t}^{(i)}\right)/\left(1+\hat{\bar{\varphi}}_x^{(i)}\hat{d}_{x,t}^{(i)}\right)}\ \forall x\\ \hat{\beta}_x^{(i+1)} = \hat{\beta}_x^{(i)}\ \forall x\\ \hat{\kappa}_t^{(i+1)} = \hat{\kappa}_t^{(i)}\ \forall t\\ \hat{\bar{\varphi}}_x^{(i+1)} = \hat{\bar{\varphi}}_x^{(i)}\ \forall x\end{cases}$

(b) Update $\{\beta_x\}$: $\begin{cases}\hat{\alpha}_x^{(i+2)} = \hat{\alpha}_x^{(i+1)}\ \forall x\\ \tilde{\beta}_x^{(i+2)} = \hat{\beta}_x^{(i+1)} - \frac{\sum_t\left(d_{x,t}\hat{\kappa}_t^{(i+1)}-\hat{\bar{\varphi}}_x^{(i+1)}\hat{\kappa}_t^{(i+1)}g_{x,t}^{(i+1)}\right)}{\sum_t\left(-\hat{\bar{\varphi}}_x^{(i+1)}\left(\hat{\kappa}_t^{(i+1)}\right)^2 g_{x,t}^{(i+1)}\right)/\left(1+\hat{\bar{\varphi}}_x^{(i+1)}\hat{d}_{x,t}^{(i+1)}\right)}\ \forall x\\ \tilde{\kappa}_t^{(i+2)} = \hat{\kappa}_t^{(i+1)}\ \forall t\\ \hat{\bar{\varphi}}_x^{(i+2)} = \hat{\bar{\varphi}}_x^{(i+1)}\ \forall x\end{cases}$

(c) Apply identification constraint (2.6): $\begin{cases}\hat{\beta}_x^{(i+2)} = \tilde{\beta}_x^{(i+2)}/\sum_x\tilde{\beta}_x^{(i+2)}\ \forall x\\ \hat{\kappa}_t^{(i+2)} = \tilde{\kappa}_t^{(i+2)}\sum_x\tilde{\beta}_x^{(i+2)}\ \forall t\end{cases}$

(d) Update $\{\kappa_t\}$: $\begin{cases}\tilde{\alpha}_x^{(i+3)} = \hat{\alpha}_x^{(i+2)}\ \forall x\\ \hat{\beta}_x^{(i+3)} = \hat{\beta}_x^{(i+2)}\ \forall x\\ \tilde{\kappa}_t^{(i+3)} = \hat{\kappa}_t^{(i+2)} - \frac{\sum_x\left(d_{x,t}\hat{\beta}_x^{(i+2)}-\hat{\bar{\varphi}}_x^{(i+2)}\hat{\beta}_x^{(i+2)}g_{x,t}^{(i+2)}\right)}{\sum_x\left(-\hat{\bar{\varphi}}_x^{(i+2)}\left(\hat{\beta}_x^{(i+2)}\right)^2 g_{x,t}^{(i+2)}\right)/\left(1+\hat{\bar{\varphi}}_x^{(i+2)}\hat{d}_{x,t}^{(i+2)}\right)}\ \forall t\\ \hat{\bar{\varphi}}_x^{(i+3)} = \hat{\bar{\varphi}}_x^{(i+2)}\ \forall x\end{cases}$

(e) Apply identification constraint (2.7): $\begin{cases}\hat{\alpha}_x^{(i+3)} = \tilde{\alpha}_x^{(i+3)} + \frac{\sum_{t=t_1}^{t_n}\tilde{\kappa}_t^{(i+3)}}{t_n-t_1+1}\hat{\beta}_x^{(i+3)}\ \forall x\\ \hat{\kappa}_t^{(i+3)} = \tilde{\kappa}_t^{(i+3)} - \frac{\sum_{t=t_1}^{t_n}\tilde{\kappa}_t^{(i+3)}}{t_n-t_1+1}\ \forall t\end{cases}$

(f) Update $\{\varphi_t\}$:

$$\begin{cases}\tilde{\alpha}_x^{(i+4)} = \hat{\alpha}_x^{(i+3)}\ \forall x\\ \hat{\beta}_x^{(i+4)} = \hat{\beta}_x^{(i+3)}\ \forall x\\ \hat{\kappa}_t^{(i+4)} = \hat{\kappa}_t^{(i+3)}\ \forall t\\ \hat{\bar{\varphi}}_x^{(i+4)} = \hat{\bar{\varphi}}_x^{(i+3)} - \frac{\sum_t\left(f_{x,t}^{(i+3)}-g_{x,t}^{(i+3)}+r_{x,t}^{(i+3)}+\frac{d_{x,t}}{\hat{\bar{\varphi}}_x^{(i+3)}}\right)}{\sum_t\left(h_{x,t}^{(i+3)}-\frac{d_{x,t}+2r_{x,t}^{(i+3)}\hat{\bar{\varphi}}_x^{(i+3)}}{\left(\hat{\bar{\varphi}}_x^{(i+3)}\right)^2}+\left(\frac{2}{\left(\hat{\bar{\varphi}}_x^{(i+3)}\right)^2}+g_{x,t}^{(i+3)}\right)\left(\frac{\hat{d}_{x,t}^{(i+3)}}{1+\hat{\bar{\varphi}}_x^{(i+3)}\hat{d}_{x,t}^{(i+3)}}\right)\right)}\ \forall x\end{cases}$$

Jump-off bias correction

The estimation procedure above results in the same jump-off bias as mentioned in section 2.3. To correct for the jump-off bias, replace the following steps in the updating scheme:

(a) Update $\{\alpha_x\}$ **once**: $\begin{cases} \hat{\alpha}_x^{(i+1)} = \ln m_{x,t_n} \ \forall x \\ \hat{\beta}_x^{(i+1)} = \hat{\beta}_x^{(i)} \ \forall x \\ \hat{\kappa}_t^{(i+1)} = \hat{\kappa}_t^{(i)} \ \forall t \end{cases}$

(e) Apply identification constraint (2.18): $\begin{cases} \hat{\alpha}_x^{(i+3)} = \tilde{\alpha}_x^{(i+3)} + \tilde{\kappa}_{t_n}^{(i+3)} \hat{\beta}_x^{(i+3)} \ \forall x \\ \hat{\kappa}_t^{(i+3)} = \tilde{\kappa}_t^{(i+3)} - \tilde{\kappa}_{t_n}^{(i+3)} \ \forall t \end{cases}$

3.4 Mortality forecasting

As in the Poisson Lee-Carter model, the Poisson-gamma Lee-Carter model reduces the difficulty of forecasting future mortality to forecasting one single mortality index κ_t. The forecast of the mortality index is obtained by finding an appropriate ARIMA(p, d, q) time series model in the same way as described in section 2.4.

Let us denote with $\hat{\kappa}_{t_n+s}$ the s-period ahead forecast of the mortality index. The expected future death count in the Poisson-gamma Lee-Carter model is given by

$$E\left[D_{x,t_n+s}\right] = e_{x,t_n+s}\hat{m}_{x,t_n+s}E[\bar{Z}_x] \tag{3.17}$$

where e_{x,t_n+s} is the future exposure and $\hat{m}_{x,t_n+s}E[\bar{Z}_x]$ is the forecast of future death rate. Intuitively $E[\bar{Z}_x]$ can be interpreted as the mean experience factor. Since we impose the assumption that $E[\bar{Z}_x] = 1$, the forecast of future death rate is given by

$$\hat{m}_{x,t_n+s}E[\bar{Z}_x] = e^{\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_{t_n+s}} \tag{3.18}$$

From the forecasted death rate we can calculate other quantities of interest, such as life expectancies, life annuity premiums etc.

3.5 Simulation and confidence intervals

To quantify the uncertainty of forecasted quantities within the Poisson-gamma Lee-Carter model, Li *et al.* (2009) extend the semi-parametric bootstrap approach proposed by Brouhns *et al.* (2005). Essentially, Li *et al.* (2009) generate bootstrap samples with the negative binomial distribution instead of the Poisson distribution.

Simulation algorithm

The extended bootstrap approach consists of the following steps:

1. Generate N bootstrap samples of the death count $d_{x,t}^{(n)}$ for $n = 1, \dots, N$, where $d_{x,t}^{(n)}$ are realizations from the negative binomial distribution given by (3.12).
2. For the nth bootstrap sample:
   (a) Re-estimate the Poisson-gamma Lee-Carter model obtaining the parameter sets $\left\{\alpha_x^{(n)}\right\}$, $\left\{\beta_x^{(n)}\right\}$, $\left\{\kappa_t^{(n)}\right\}$, $\left\{\bar{\varphi}_x^{(n)}\right\}$;

   (b) Using $\left\{\kappa_t^{(n)}\right\}$ re-estimate the parameters of the originally estimated ARIMA model where the order of integration d, the degree p of the autoregressive polynomial and the degree q of the moving average polynomial do not change;
   (c) Generate a projection of future mortality index $\kappa_t^{(n)}$ using the ARIMA model obtained in step (2b);
   (d) Calculate the forecasted future mortality rate;
   (e) Calculate the quantity of interest using the forecasted future mortality rate.
3. Derive the confidence interval for the quantity of interest using the empirical distribution obtained in step (2).

Chapter 4

The Lee-Carter model in a Poisson-gamma setting with general dispersion parameter

Chapter 3 discussed the Poisson-gamma Lee-Carter model with age-specific dispersion parameters introduced by Li *et al.* (2009). This chapter describes the approach with a general dispersion parameter considered by Delwarde *et al.* (2007), which we use for comparison in section 5.3 to test whether the generalization to age-specific dispersion parameters leads to a significant improvement of the fit.

Section 4.1 presents the specification of the Poisson-gamma Lee-Carter model with general dispersion parameter. Section 4.2 contains the numerical methods to obtain the corresponding maximum-likelihood. We do not elaborate on the subject of jump-off bias correction, mortality forecasting, simulation and confidence intervals, since the techniques discussed in chapter 3 also apply to the approach in this chapter albeit with some minor modifications.

4.1 The model specification

The Poisson-gamma Lee-Carter model with general dispersion parameter equals the Poisson-gamma Lee-Carter model given by (3.11) with the main difference being the assumption that the structural variable has a distribution with non-age specific parameters. Applying this assumption leads to the following variant of the Poisson-gamma Lee-Carter model

$$\begin{cases} D_{x,t}|\bar{Z} \sim \text{Poisson}\left(e_{x,t} m_{x,t} \bar{Z}\right) \\ \bar{Z} \sim Gamma(\bar{\varphi}^{-1}, \bar{\varphi}^{-1}) \\ \text{with } m_{x,t} = e^{\alpha_x + \beta_x \kappa_t} \end{cases} \tag{4.1}$$

which has unconditional mean and variance

$$E\left[D_{x,t}\right] = e_{x,t} m_{x,t} \tag{4.2}$$

$$Var\left(D_{x,t}\right) = e_{x,t} m_{x,t} + \bar{\varphi}\left(e_{x,t} m_{x,t}\right)^2 \tag{4.3}$$

Similar to the Poisson-gamma model stated by (3.11), the model above retains the appealing feature that the mean equals the mean of the Poisson Lee-Carter model for

equal $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\kappa}$. Additionally this model allows for overdispersion as well, but the model uses a general dispersion parameter instead of age-specific dispersion parameters.

4.2 Estimation by maximum likelihood

To estimate the Lee-Carter model as given by (4.1), Delwarde *et al.* (2007) use the maximum likelihood method which requires the maximization of the following associated log-likelihood function

$$l(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa},\bar{\varphi};\boldsymbol{d}) = \log\prod_{x,t} f\left(d_{x,t};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa},\bar{\varphi}\right)$$

$$= \log\prod_{x,t}\binom{\bar{\varphi}^{-1}+d_{x,t}-1}{d_{x,t}}\left(\frac{1}{1+e_{x,t}m_{x,t}\bar{\varphi}}\right)^{\bar{\varphi}^{-1}}\left(1-\frac{1}{1+e_{x,t}m_{x,t}\bar{\varphi}}\right)^{d_{x,t}}$$

$$= \sum_{x,t}\left\{\sum_{i=0}^{d_{x,t}-1}\log(\bar{\varphi}^{-1}+i) - \log\left(d_{x,t}!\right) + d_{x,t}\log\left(e_{x,t}m_{x,t}\bar{\varphi}\right)\right.$$

$$\left. - \left(d_{x,t}+\bar{\varphi}^{-1}\right)\log\left(1+e_{x,t}m_{x,t}\bar{\varphi}\right)\right\} \tag{4.4}$$

Multidimensional maximization with Newton's method

Maximizing the log-likelihood function takes place using the algorithm of Goodman (1979) as described in section 2.3. Denoting $\hat{\alpha}_x^{(i)}$, $\hat{\beta}_x^{(i)}$, $\hat{\kappa}_x^{(i)}$ and $\hat{\bar{\varphi}}^{(i)}$ as the estimates at iteration *i*, we first define the following functions:

- $\hat{d}_{x,t}^{(i)} = e_{x,t}\hat{m}_{x,t} = e_{x,t}e^{\hat{\alpha}_x^{(i)}+\hat{\beta}_x^{(i)}\hat{\kappa}_x^{(i)}}$;
- $f_{x,t}^{(i)} = \sum_{i=0}^{d_{x,t}-1}\left(\frac{i}{1+\hat{\bar{\varphi}}^{(i)}\cdot i} - \frac{1}{\hat{\bar{\varphi}}^{(i)}}\right)$;
- $g_{x,t}^{(i)} = \hat{d}_{x,t}^{(i)}\left(d_{x,t}+\frac{1}{\hat{\bar{\varphi}}^{(i)}}\right)\left(1+\hat{\bar{\varphi}}^{(i)}\hat{d}_{x,t}^{(i)}\right)$;
- $h_{x,t}^{(i)} = \sum_{i=0}^{d_{x,t}-1}\left(\frac{-i^2}{\left(1+\hat{\bar{\varphi}}^{(i)}\cdot i\right)^2} - \frac{1}{\left(\hat{\bar{\varphi}}^{(i)}\right)^2}\right)$;
- $r_{x,t}^{(i)} = \frac{\log\left(1+\hat{\bar{\varphi}}^{(i)}\hat{d}_{x,t}^{(i)}\right)}{\left(\hat{\bar{\varphi}}^{(i)}\right)^2}$.

Next, the iterative scheme consists of the following updating steps:

(a) Update $\{\alpha_x\}$: $\begin{cases}\hat{\alpha}_x^{(i+1)} = \hat{\alpha}_x^{(i)} - \frac{\sum_t\left(d_{x,t}-\hat{\bar{\varphi}}^{(i)}g_{x,t}^{(i)}\right)}{-\sum_t\left(-\hat{\bar{\varphi}}^{(i)}g_{x,t}^{(i)}\right)/\left(1+\hat{\bar{\varphi}}^{(i)}\hat{d}_{x,t}^{(i)}\right)}\ \forall x\\ \hat{\beta}_x^{(i+1)} = \hat{\beta}_x^{(i)}\ \forall x\\ \hat{\kappa}_t^{(i+1)} = \hat{\kappa}_t^{(i)}\ \forall t\\ \hat{\bar{\varphi}}^{(i+1)} = \hat{\bar{\varphi}}^{(i)}\ \forall x\end{cases}$

(b) Update $\{\beta_x\}$:
$$\begin{cases} \hat{\alpha}_x^{(i+2)} = \hat{\alpha}_x^{(i+1)} \ \forall x \\ \tilde{\beta}_x^{(i+2)} = \hat{\beta}_x^{(i+1)} - \dfrac{\Sigma_t\left(d_{x,t}\hat{\kappa}_t^{(i+1)} - \hat{\varphi}^{(i+1)}\hat{\kappa}_t^{(i+1)} g_{x,t}^{(i+1)}\right)}{\Sigma_t\left(-\hat{\varphi}^{(i+1)}\left(\hat{\kappa}_t^{(i+1)}\right)^2 g_{x,t}^{(i+1)}\right)\Big/\left(1+\hat{\varphi}^{(i+1)}\hat{d}_{x,t}^{(i+1)}\right)} \ \forall x \\ \tilde{\kappa}_t^{(i+2)} = \hat{\kappa}_t^{(i+1)} \ \forall t \\ \hat{\varphi}^{(i+2)} = \hat{\varphi}^{(i+1)} \ \forall x \end{cases}$$

(c) Apply identification constraint (2.6):
$$\begin{cases} \hat{\beta}_x^{(i+2)} = \tilde{\beta}_x^{(i+2)} / \sum_x \tilde{\beta}_x^{(i+2)} \ \forall x \\ \hat{\kappa}_t^{(i+2)} = \tilde{\kappa}_t^{(i+2)} \sum_x \tilde{\beta}_x^{(i+2)} \ \forall t \end{cases}$$

(d) Update $\{\kappa_t\}$:
$$\begin{cases} \tilde{\alpha}_x^{(i+3)} = \hat{\alpha}_x^{(i+2)} \ \forall x \\ \hat{\beta}_x^{(i+3)} = \hat{\beta}_x^{(i+2)} \ \forall x \\ \tilde{\kappa}_t^{(i+3)} = \hat{\kappa}_t^{(i+2)} - \dfrac{\Sigma_x\left(d_{x,t}\hat{\beta}_x^{(i+2)} - \hat{\varphi}^{(i+2)}\hat{\beta}_x^{(i+2)} g_{x,t}^{(i+2)}\right)}{\Sigma_x\left(-\hat{\varphi}^{(i+2)}\left(\hat{\beta}_x^{(i+2)}\right)^2 g_{x,t}^{(i+2)}\right)\Big/\left(1+\hat{\varphi}^{(i+2)}\hat{d}_{x,t}^{(i+2)}\right)} \ \forall t \\ \hat{\varphi}^{(i+3)} = \hat{\varphi}^{(i+2)} \ \forall x \end{cases}$$

(e) Apply identification constraint (2.7):
$$\begin{cases} \hat{\alpha}_x^{(i+3)} = \tilde{\alpha}_x^{(i+3)} + \dfrac{\sum_{t=t_1}^{t_n} \tilde{\kappa}_t^{(i+3)}}{t_n - t_1 + 1} \hat{\beta}_x^{(i+3)} \ \forall x \\ \hat{\kappa}_t^{(i+3)} = \tilde{\kappa}_t^{(i+3)} - \dfrac{\sum_{t=t_1}^{t_n} \tilde{\kappa}_t^{(i+3)}}{t_n - t_1 + 1} \ \forall t \end{cases}$$

(f) Update $\{\varphi_t\}$:
$$\begin{cases} \tilde{\alpha}_x^{(i+4)} = \hat{\alpha}_x^{(i+3)} \ \forall x \\ \hat{\beta}_x^{(i+4)} = \hat{\beta}_x^{(i+3)} \ \forall x \\ \hat{\kappa}_t^{(i+4)} = \hat{\kappa}_t^{(i+3)} \ \forall t \\ \hat{\varphi}^{(i+4)} = \hat{\varphi}^{(i+3)} - \dfrac{\Sigma_{x,t}\left(f_{x,t}^{(i+3)} - g_{x,t}^{(i+3)} + r_{x,t}^{(i+3)} + \frac{d_{x,t}}{\hat{\varphi}^{(i+3)}}\right)}{\Sigma_{\mathrm{x,t}}\left(h_{x,t}^{(i+3)} - \frac{d_{x,t} + 2r_{x,t}^{(i+3)}\hat{\varphi}^{(i+3)}}{\left(\hat{\varphi}^{(i+3)}\right)^2} + \left(\frac{2}{\left(\hat{\varphi}^{(i+3)}\right)^2} + g_{x,t}^{(i+3)}\right)\left(\frac{\hat{d}_{x,t}^{(i+3)}}{1+\hat{\varphi}^{(i+3)}\hat{d}_{x,t}^{(i+3)}}\right)\right)} \ \forall x \end{cases}$$

Chapter 5

An application to Dutch population mortality data

In this chapter we compare the performance of the Poisson approach of the Lee-Carter model to both Poisson-gamma approaches. For this purpose we use the Dutch mortality data obtained from the Human Mortality Database which is described in section 5.1. Section 5.2 contains the parameter estimates of the three models. In section 5.3 we perform a comparison by analyzing the quality of the fit. Section 5.4 contains the estimated ARIMA-models used for forecasting the mortality index. Note that for mortality forecasting we do not take the Poisson-gamma approach with general dispersion parameter into account. Section 5.5 presents the results of the Dutch mortality forecast. In section 5.6 we assess the forecasting abilities of the two models using backtesting. Section 5.7 compares the forecasts of the two models with the projection of the mortality rate for 2010 to 2060 as published by the Dutch Actuarial Association.

5.1 Description of the data

The Human Mortality Database (HMD) provides national mortality data of diverse countries[1]. In our thesis we use the HMD as our main source for the Dutch population mortality data. The datasets provided by HMD include the number of deaths, the exposure and death rates available in the following formats of age and time:

- Age groups of one-year or five-year;
- Time intervals of one-year, five-year or ten-year.

The datasets cover the ages 0, 1, …, 109, 110+ where the latter refers to the age group containing the ages 110 and older. At the time of writing the historical data are available for the period from 1850 to 2009. The HMD bases its mortality data from 1950 and later on the official data on births, deaths, and population provided by Statistics Netherlands. The mortality data for the period from 1850 to 1949 come from the NIDI mortality database.

For our purposes we exclude the years prior to 1950 hereby preventing the effects of the Spanish flu in 1918, the First and Second World War affecting the parameter

[1] http://www.mortality.org

estimates. Furthermore, we exclude the age groups of 100 and higher, because these age groups contain low numbers of observations.

5.2 Parameter estimates of the Lee-Carter model

We obtain the Lee-Carter parameter estimates by fitting the Poisson and both Poisson-gamma models to the Dutch mortality data from 1950 to 2009. Note that we choose to apply the jump-off correction to the fitting procedure. Figure 5.1 to Figure 5.3 depict the obtained Lee-Carter parameter estimates $\hat{\alpha}_x$, $\hat{\beta}_x$, $\hat{\kappa}_t$. Figure 5.4 shows the estimated dispersion parameters $\hat{\bar{\varphi}}$ and $\hat{\bar{\varphi}}_x$.

The estimated $\hat{\alpha}_x$'s of the three models are identical, since the jump-off correction imposes the constraint $\hat{\alpha}_x = \ln m_{x,t_n}$. For the estimated $\hat{\beta}_x$, both Poisson-gamma models show no visible difference with the Poisson model. Concerning the estimated $\hat{\kappa}_t$, the three models show no visible difference for females. However, for males we observe a small difference, but the overall long-term behavior of the mortality index seems to remain unaffected.

Next, we analyze the estimated dispersion parameters $\hat{\bar{\varphi}}_x$ and $\hat{\bar{\varphi}}$ which only apply to the Poisson-gamma models. A positive value for $\hat{\bar{\varphi}}$ indicates the general presence of overdispersion over all ages, while a positive $\hat{\bar{\varphi}}_x$ indicates the presence of overdispersion for age x. For males we observe peaks for $\hat{\bar{\varphi}}_x$ between the age of 0 and 20, while for females we see peaks between the age of 15 to 35. The observed peaks for males are relatively higher than for females. Furthermore the general dispersion parameter $\hat{\bar{\varphi}}$ for males lies higher than for females. From this we conclude that the male mortality data contain more overdispersion than the female mortality data. Since the Poisson-gamma models capture more variability than the Poisson model, this indicates that the Poisson-gamma models are relatively more effective for males than for females.

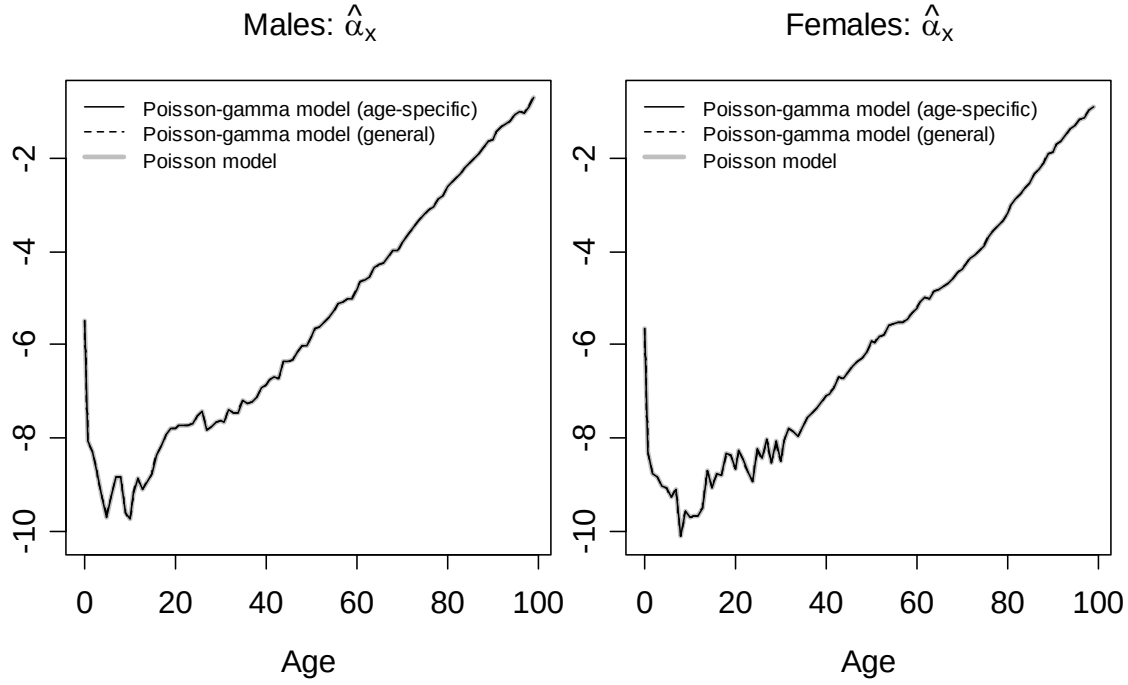

Figure 5.1: Estimates of parameter α_x

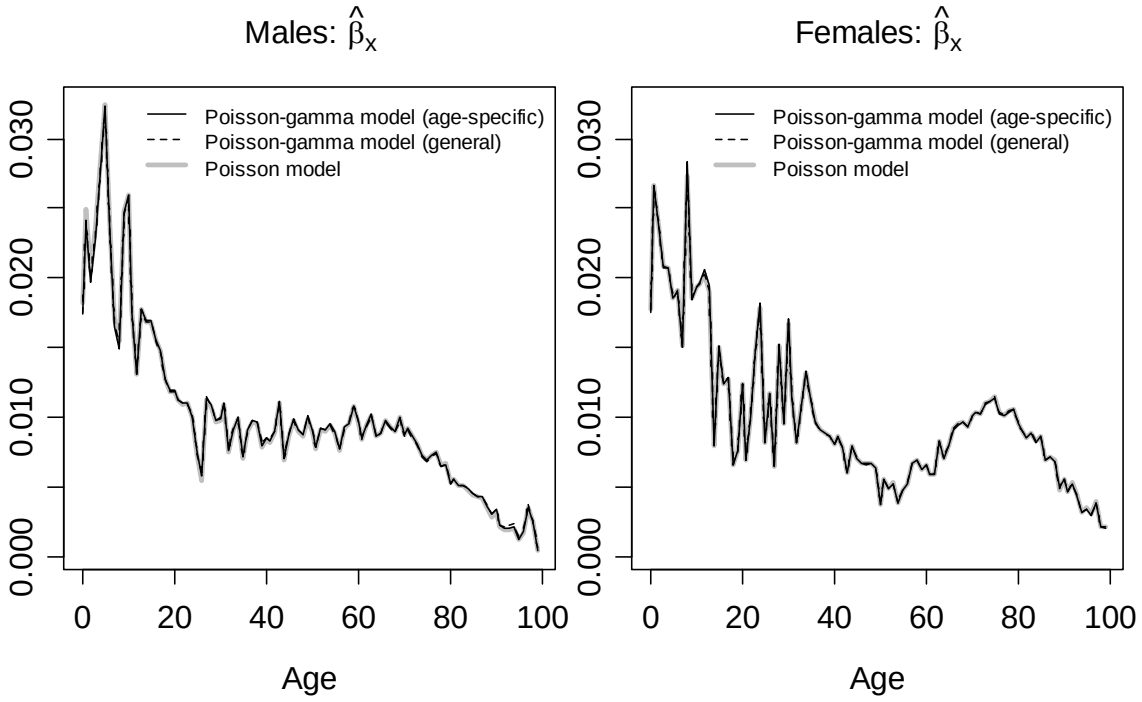

Figure 5.2: Estimates of parameter β_x

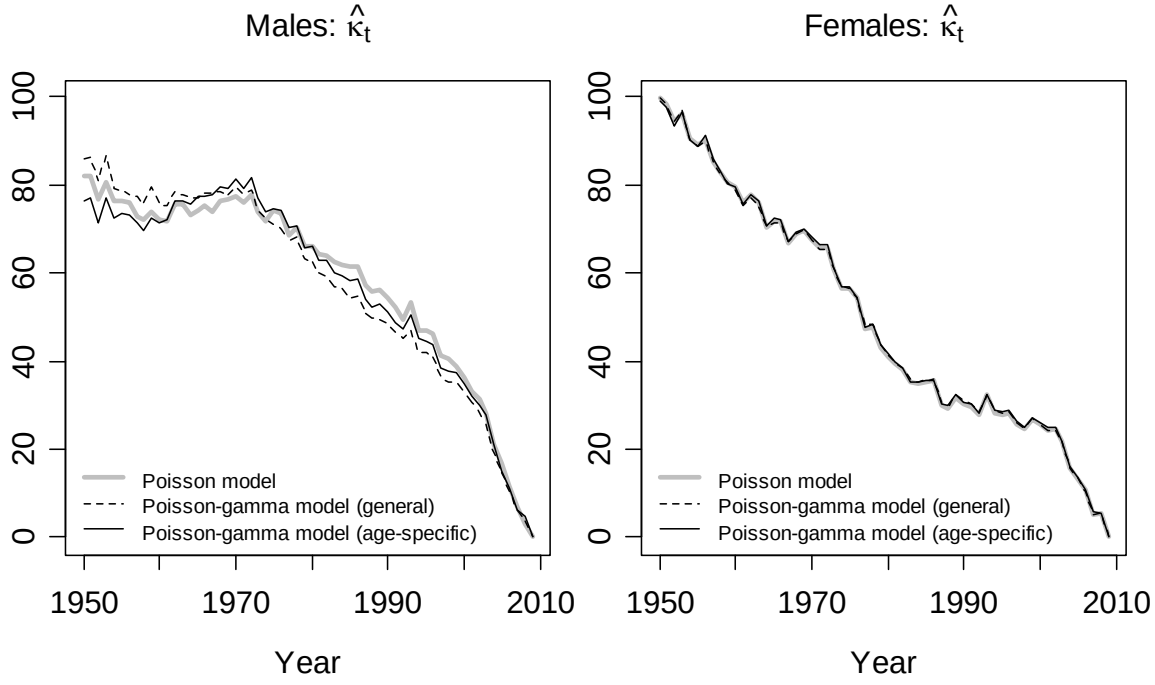

Figure 5.3: Estimates of parameter κ_t

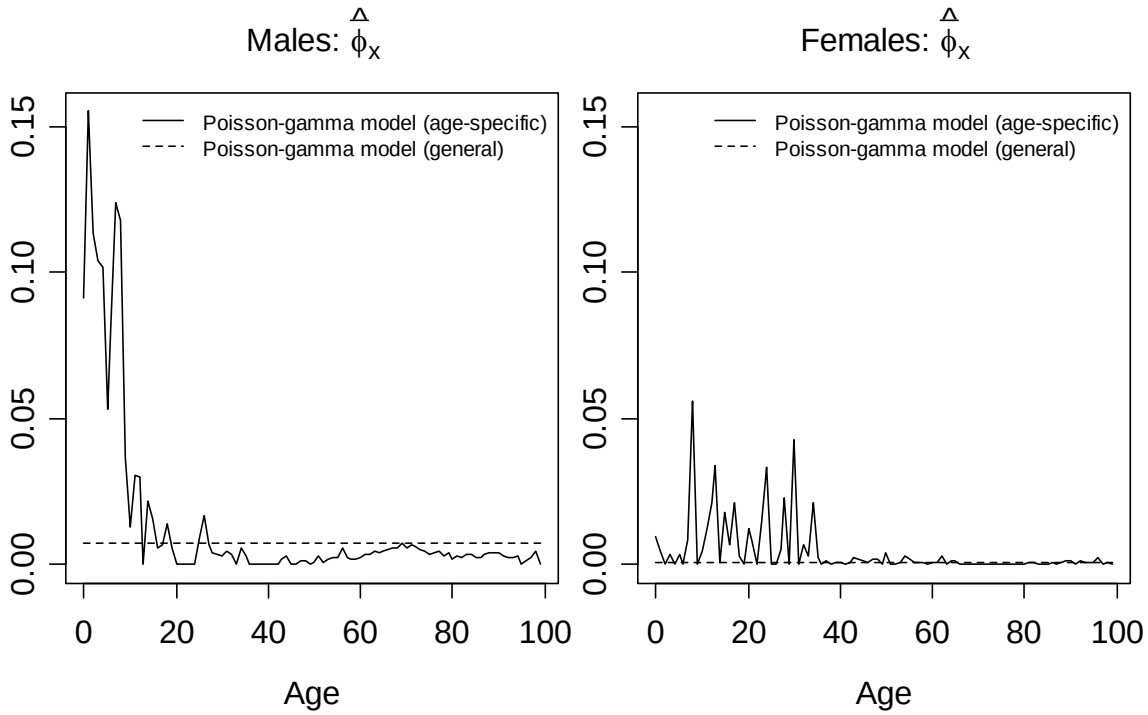

Figure 5.4: Estimates of the dispersion parameters $\bar{\varphi}$ and $\bar{\varphi}_x$

5.3 Quality of the fit

In this section we analyze the quality of fit and compare the three Lee-Carter models. To this end we use information criterion measures, tests for detecting overdispersion and the Likelihood-Ratio-Test.

Information criterion measures

The Poisson-gamma model with age-specific dispersion parameters gives the highest log-likelihood compared to the other two models as shown in Table 5.1. However, since this model also uses more parameters, we cannot conclude its superiority with respect to the other two models. To compare the three models we use the following information criterion:

- Akaike Information Criterion (AIC) given by $AIC = -2\ln(L) + 2k$;
- Bayesian Information Criterion (BIC) given by $BIC = -2\ln(L) + k\ln(n)$.

where L denotes the maximized value of the likelihood function, k the number of parameters and n the number of observations. The information criterion above not only looks at the likelihood, but also includes a penalty as an increasing function of the number of estimated parameters to discourage overfitting. Thus the model with the lowest information criterion has our preference.

The AIC and BIC in Table 5.1 indicate for males that the Poisson-gamma model with age-specific dispersion parameters results in the best fit compared to the other two models. For females both Poisson-gamma models lead to a better fit than the Poisson model. However, the AIC and BIC do not give a unanimous conclusion concerning the superiority (or inferiority) of age-specific dispersion parameters over a general dispersion parameter. In comparison with the Poisson model, both Poisson-gamma models improve

Model:	Poisson		Poisson-gamma (general)		Poisson-gamma (age-specific)	
	Males	**Females**	**Males**	**Females**	**Males**	**Females**
Number of parameters	260	260	261	261	360	360
Log-likelihood	-36,967	-25,696	-29,134	-25,507	-28,228	-25,234
AIC	74,454	51,911	58,790	51,536	57,176	51,188
BIC	76,196	53,653	60,538	53,285	59,588	53,600

Table 5.1: Summary statistics for the Poisson model and both Poisson-gamma models

the fit relatively more for males than for females, which conforms to our (visual) observations of $\hat{\tilde{\varphi}}_x$ and $\hat{\tilde{\varphi}}$ in section 5.2.

Overdispersion tests

To detect overdispersion Denuit *et al.* (2007) propose several test statistics for testing a Poisson model against heterogeneity models with a variance function of the form

$$Var(N_i) = E[N_i] + \tau(E[N_i])^2 \tag{5.1}$$

with $\tau = Var(\Theta_i)$ being the variance of the random effect. More specifically, we need to test the null hypothesis $H_0: \tau = 0$ against $H_1: \tau > 0$.

The Poisson-gamma Lee-Carter model with general dispersion parameter has a variance function given by equation (4.3) which has form (5.1). Therefore the three test statistics described in Denuit *et al.* (2007) can be adapted to our case as follows

$$T_1 = \frac{\sum_{x,t}\left(\left(d_{x,t} - \hat{d}_{x,t}\right)^2 - d_{x,t}\right)}{\sqrt{2\sum_{x,t}\hat{d}_{x,t}^2}} \sim N(0,1) \tag{5.2}$$

$$T_2 = \frac{\sum_{x,t}\left(\left(d_{x,t} - \hat{d}_{x,t}\right)^2 - d_{x,t}\right)}{\sqrt{\sum_{x,t}\left(\left(d_{x,t} - \hat{d}_{x,t}\right)^2 - d_{x,t}\right)^2}} \sim N(0,1) \tag{5.3}$$

$$T_3 = \frac{\sum_{x,t}\left(\left(d_{x,t} - \hat{d}_{x,t}\right)^2 - d_{x,t}\right)}{\sqrt{\frac{1}{n}\sum_{x,t}\hat{d}_{x,t}^{-2}\left(\left(d_{x,t} - \hat{d}_{x,t}\right)^2 - d_{x,t}\right)^2}\sqrt{\sum_{x,t}\hat{d}_{x,t}^2}} \sim N(0,1) \tag{5.4}$$

where $\hat{d}_{x,t} = e_{x,t}e^{\hat{\alpha}_x + \hat{\beta}_x\hat{\kappa}_t}$ denotes the estimated death count and n denotes the number of observations. Table 5.2 contains the values of the test statistics for males and females. All associated p-values are less than 10^{-6} and for that reason we reject the null hypothesis in favor of the Poisson-gamma Lee-Carter model with general dispersion parameter for both males and females.

	Males	**Females**
T_1	287.40	21.85
T_2	9.54	9.97
T_3	20.22	12.56

Table 5.2: Values of the test statistics T_1, T_2 and T_3

Likelihood-Ratio-Test

If one model (the null model) forms a submodel of the other (the full model) and the log-likelihood of both models are available, then the Likelihood-Ratio-Test (LRT) can be applied to compare the goodness-of-fit of the two models. Under the null hypothesis that the null model is correct the LRT defines the following test statistic

$$D = -2\ln\left(\frac{L_{null}}{L_{full}}\right) = -2\big(\ln L_{null} - \ln L_{full}\big) \sim \chi^2(r) \tag{5.5}$$

where L_{null} and L_{full} denote the likelihood of the null and full model respectively. The χ^2-distribution has r degrees of freedom equal to the number of additional parameters.

The LRT is applied by Li *et al.* (2009) to test whether the Poisson-gamma Lee-Carter model with age-specific dispersion parameters leads to a significant improvement over the Poisson Lee-Carter model. However, such a test is incorrect, because the null hypothesis lies on the boundary of the parameter space of the Poisson-gamma. As a consequence, the test statistic D has no longer an asymptotic χ2-distribution.

We use the LRT to compare the fit between the two types of Poisson-gamma models, i.e. comparison between the age-specific dispersion parameter and the general dispersion parameter. For males $D = 1{,}812$ and for females $D = 546$, which have corresponding p-values less than 10^{-6}. Therefore, together with the results of the previous subsection, we conclude that the Poisson-gamma Lee-Carter model with age-specific dispersion parameters gives the significantly best fit for both males and females.

5.4 ARIMA-model for the mortality index

For the Lee-Carter models estimated in section 5.2 we fit an appropriate ARIMA-model for the mortality index κ_t which we later use to forecast future mortality rates. We follow the Box-Jenkins methodology as described in section 2.4. Note that we do not take the Poisson-gamma Lee-Carter model with general dispersion parameter into account.

In Figure 5.3 we observe that the κ_t's for females exhibit a clear downward linear trend which indicates an ARIMA-model with integration order of one. The κ_t's for males show a slightly downward quadratic trend indicating an ARIMA-model with integration order equal to two. Since the $\{\kappa_t\}$ coefficients reflect the general level of mortality in time (the mortality index), we prefer the ARIMA-model for males to have the same order of integration as the ARIMA-model for females. At the end of this section we address second order integrated ARIMA-model in more detail.

To determine the orders p and q of the autoregressive and respectively moving average polynomial one usually starts by comparing the sample ACF and sample PACF with their theoretical values. However, for an ARIMA-model with a combination of autoregressive and moving average properties it is difficult to determine the order of the polynomials merely from the sample ACF and sample PACF. Therefore we look at several ARIMA-models which pass the diagnostic checking, i.e. ARIMA-models with uncorrelated residuals. To test the absence of autocorrelation we apply the Ljung-Box Q-test as given by (2.29). Additionally we calculate the AIC and BIC for these models as given by (2.27) and (2.28) respectively. Finally we choose the most parsimonious ARIMA-model indicated by the lowest information criterion value.

Results for the male mortality index of the Poisson Lee-Carter model

Table 5.3 shows the first order integrated ARIMA-models fitted to the male mortality index κ_t from the Poisson Lee-Carter model, which pass the diagnostic checking. From the values of AIC and BIC we conclude ARIMA(0, 1, 3) to be the most appropriate model. The other ARIMA-models show higher values of AIC and BIC, or even have insignificant coefficients. Figure 5.5 shows the plot of the Ljung-Box Q-test performed on the residuals of the ARIMA(0, 1, 3)-model. The p-values lie well above the 5% level of significance, which means that the null hypothesis of no autocorrelation stands true. Figure 5.9 depicts the mean forecast together with the 95% confidence interval.

ARIMA	**(4, 1, 0)**	**(1, 1, 2)**	**(2, 1, 2)**	**(0, 1, 3)**	**(1, 1, 3)**	**(0, 1, 4)**
c	-1.7632	-1.4365	-1.8784	-1.3997	-1.3997	-1.3997
(s.e.)	*(0.9049)*	*(0.4774)*	*(1.0466)*	*(0.4316)*	*(0.4313)*	*(0.4312)*
φ_1	-0.3329	0.4308	0.8689	-1.3997	-0.0015	
(s.e.)	*(0.1200)*	*(0.1282)*	*(0.2980)*	(0.4316)	*(0.2327)*	
φ_2	0.1610		0.0661			
(s.e.)	*(0.1033)*		*(0.2752)*			
φ_3	0.5463					
(s.e.)	*(0.1100)*					
φ_4	0.3729					
(s.e.)	*(0.1288)*					
θ_1		-0.8414	-1.2122	-0.2991	-0.2980	-0.2996
(s.e.)		*(0.0704)*	*(0.2907)*	*(0.1122)*	*(0.2004)*	*(0.1340)*
θ_2		0.8618	0.4881	0.4586	0.4579	0.4583
(s.e.)		*(0.1329)*	*(0.2261)*	*(0.1152)*	*(0.1630)*	*(0.1258)*
θ_3				0.5287	0.5295	0.5288
(s.e.)				*(0.1290)*	*(0.1820)*	*(0.1311)*
θ_4						-0.0010
(s.e.)						*(0.1358)*
AIC	264.42	267.63	269.78	263.12	265.12	265.12
BIC	276.89	278.02	282.25	273.5	277.58	277.58

Table 5.3: ARIMA-models fitted to κ_t from the Poisson Lee-Carter model for males

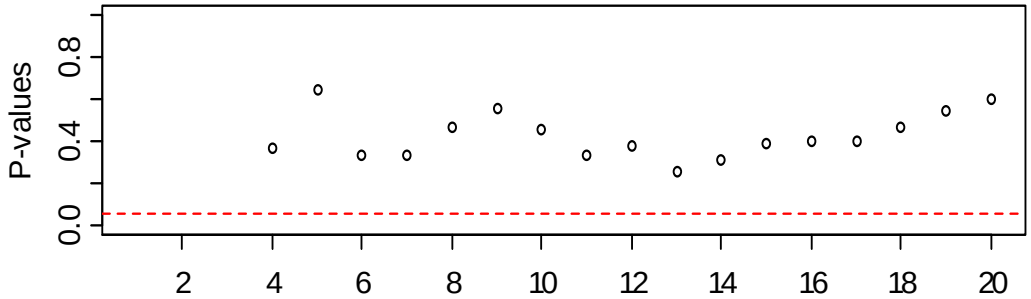

Figure 5.5: Ljung-Box Q-test performed on the residuals of the ARIMA(0, 1, 3)-model fitted to κ_t from the Poisson Lee-Carter model for males

Results for the male mortality index of the Poisson-gamma Lee-Carter model

Table 5.4 contains the results of the ARIMA-models with integration order of one fitted to the male mortality index κ_t from the Poisson-gamma Lee-Carter model, which pass the diagnostic checking. Similar to the results of Poisson Lee-Carter model we conclude again that ARIMA(0, 1, 3) is the most appropriate model based on the lower values of AIC and BIC. Figure 5.6 depicts the corresponding result of the Ljung-Box *Q*-test from which we conclude that the residuals has no autocorrelation. The forecast results are shown in Figure 5.9.

ARIMA	**(4, 1, 0)**	**(1, 1, 2)**	**(0, 1, 3)**	**(1, 1, 3)**	**(0, 1, 4)**
c	-1.4856	-1.3024	-1.2891	-1.2962	-1.2999
(s.e.)	*(0.7940)*	*(0.5127)*	*(0.4224)*	*(0.4754)*	*(0.4767)*
φ_1	-0.3059	0.4936		0.2292	
(s.e.)	*(0.1227)*	*(0.1302)*		*(0.2451)*	
φ_2	0.2381				
(s.e.)	*(0.1150)*				
φ_3	0.4336				
(s.e.)	*(0.1245)*				
φ_4	0.3106				
(s.e.)	*(0.1356)*				
θ_1		-0.8528	-0.3524	-0.5299	-0.2857
(s.e.)		*(0.0967)*	*(0.1093)*	*(0.2318)*	*(0.1327)*
θ_2		0.7742	0.4257	0.5424	0.4771
(s.e.)		*(0.1479)*	*(0.1013)*	*(0.1793)*	*(0.1201)*
θ_3			0.4299	0.3055	0.3916
(s.e.)			*(0.1152)*	*(0.1956)*	*(0.1234)*
θ_4					0.1344
(s.e.)					*(0.1318)*
AIC	274.19	273.73	272.71	273.91	273.68
BIC	286.66	284.12	283.09	286.38	286.14

Table 5.4: ARIMA-models fitted to κ_t from the Poisson-gamma Lee-Carter model for males

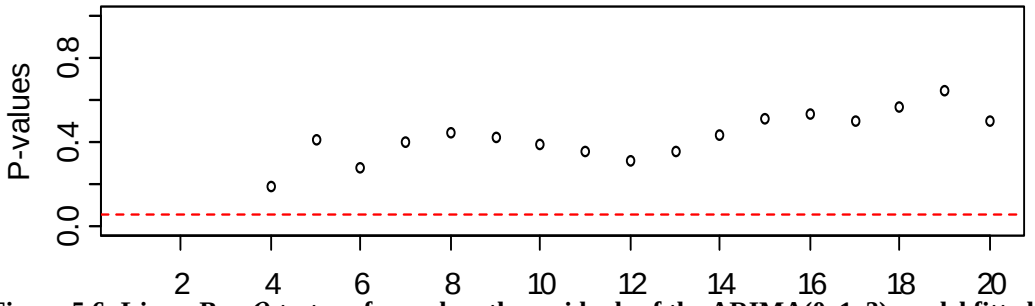

Figure 5.6: Ljung-Box *Q*-test performed on the residuals of the ARIMA(0, 1, 3)-model fitted to κ_t from the Poisson-gamma Lee-Carter model for males

Results for the female mortality index of the Poisson Lee-Carter model

The ARIMA-models fitted to the female mortality index of the Poisson Lee-Carter model, which pass the diagnostic checking, are shown in Table 5.5. The most parsimonious ARIMA-model seems to be ARIMA(0, 1, 3) according to the values of AIC and BIC. Other models have a higher AIC-value and BIC-value, or have insignificant coefficients. Therefore we conclude ARIMA(0, 1, 3) to be the most appropriate model. Figure 5.7 depicts the corresponding result of the Ljung-Box Q-test which indicates that the residuals have no autocorrelation. The forecast results are depicted in Figure 5.9.

ARIMA	**(3, 1, 0)**	**(4, 1, 0)**	**(3, 1, 1)**	**(3, 1, 2)**	**(0, 1, 3)**	**(1, 1, 3)**	**(0, 1, 4)**
c	-1.6976	-1.7602	-1.7462	-1.7143	-1.7017	-1.6987	-1.697
(s.e.)	*(0.3115)*	*(0.4078)*	*(0.3908)*	*(0.3496)*	*(0.3601)*	*(0.3364)*	*(0.319)*
φ_1	-0.2301	-0.3368	0.0635	0.1693		-0.1638	
(s.e.)	*(0.1224)*	*(0.1264)*	*(0.1987)*	*(0.2613)*		*(0.2564)*	
φ_2	-0.0564	-0.0282	0.0391	-0.2063			
(s.e.)	*(0.1245)*	*(0.1181)*	*(0.1253)*	*(0.1986)*			
φ_3	0.3743	0.4247	0.4409	0.3659			
(s.e.)	*(0.1249)*	*(0.1212)*	*(0.1195)*	*(0.1527)*			
φ_4		0.2906					
(s.e.)		*(0.1319)*					
θ_1			-0.3474	-0.4930	-0.3089	-0.1843	-0.3726
(s.e.)			*(0.1976)*	*(0.2909)*	*(0.1208)*	*(0.2207)*	*(0.1389)*
θ_2				0.3611	0.1996	0.1373	0.1890
(s.e.)				*(0.2510)*	*(0.1581)*	*(0.1710)*	*(0.1403)*
θ_3					0.4208	0.4756	0.4575
(s.e.)					*(0.1415)*	*(0.1561)*	*(0.1440)*
θ_4							-0.1089
(s.e.)							*(0.1332)*
AIC	271.57	268.97	271.24	270.97	268.5	270.09	269.84
BIC	281.96	281.43	283.71	285.51	278.89	282.55	282.3

Table 5.5: ARIMA-models fitted to κ_t from the Poisson Lee-Carter model for females

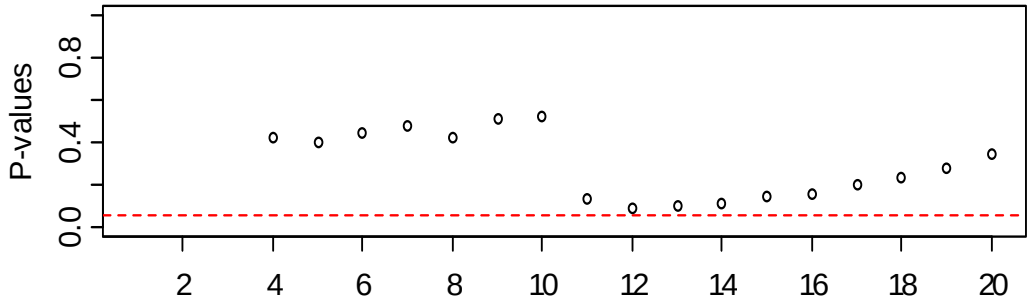

Figure 5.7: Ljung-Box Q-test performed on the residuals of the ARIMA(0, 1, 3)-model fitted to κ_t from the Poisson Lee-Carter model for females

Results for the female mortality index of the Poisson-gamma Lee-Carter model

The results of the estimated ARIMA-models, which pass the diagnostic checking, for the female mortality index of the Poisson-gamma Lee-Carter model are summarized in Table 5.6. ARIMA(4, 1, 0) and ARIMA(2, 1, 2) have the two highest value of BIC and AIC. Nonetheless for comparison we prefer to keep the same ARIMA(0, 1, 3)-model as for the Poisson Lee-Carter model, despite the second AR-coefficient being not significantly different from zero. Figure 5.8 shows the result of the Ljung-Box *Q*-test performed on the residuals of the ARIMA(0, 1, 3)-model. We conclude that the null hypothesis of no autocorrelation stands true, albeit some *p*-values lie very close to the 5% level of significance. Figure 5.9 depicts the results of the forecast.

ARIMA	**(3, 1, 0)**	**(4, 1, 0)**	**(3, 1, 1)**	**(2, 1, 2)**	**(3, 1, 2)**	**(2, 1, 3)**	**(0, 1, 3)**
c	-1.6813	-1.7348	-1.7238	-1.6648	-1.6979	-1.6827	-1.6880
(s.e.)	*(0.3163)*	*(0.4039)*	*(0.3905)*	*(0.2419)*	*(0.3519)*	*(0.3018)*	*(0.3441)*
φ_1	-0.2394	-0.3359	0.0457	-0.5890	0.0888	-0.1258	
(s.e.)	*(0.1222)*	*(0.1272)*	*(0.2070)*	*(0.1216)*	*(0.2708)*	*(0.3415)*	
φ_2	-0.0616	-0.0365	0.0337	-0.8624	-0.1706	-0.2106	
(s.e.)	*(0.1246)*	*(0.1193)*	*(0.1266)*	*(0.0857)*	*(0.2145)*	*(0.2444)*	
φ_3	0.3782	0.4257	0.4415		0.3826		
(s.e.)	*(0.1249)*	*0.1224)*	*(0.1192)*		*(0.1414)*		
φ_4		0.2655					
(s.e.)		*(0.1337)*					
θ_1			-0.3364	0.3010	-0.4001	-0.2173	-0.2967
(s.e.)			*(0.2088)*	*(0.1527)*	*(0.2978)*	*(0.3308)*	*(0.1251)*
θ_2				0.7699	0.2748	0.2673	0.1236
(s.e.)				*(0.1330)*	*(0.2126)*	*(0.3067)*	*(0.1566)*
θ_3						0.3537	0.3645
(s.e.)						*(0.1731)*	*(0.1263)*
AIC	274.75	272.98	274.75	273.32	275.26	276.7	273.8
BIC	285.14	285.45	287.21	285.79	289.8	291.24	284.19

Table 5.6: ARIMA-models fitted to κ_t from the Poisson-gamma Lee-Carter model for females

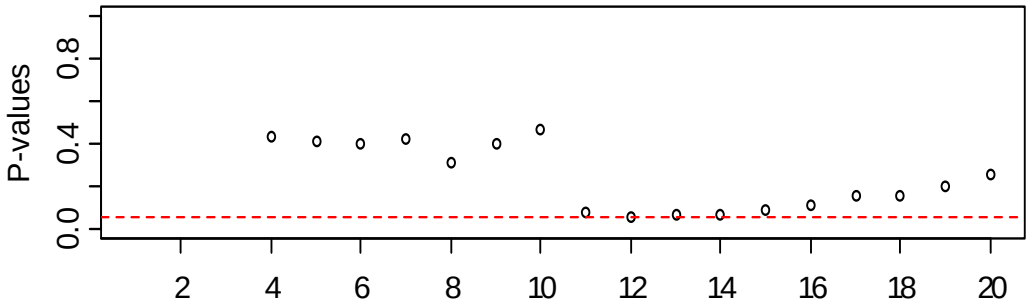

Figure 5.8: Ljung-Box *Q*-test performed on the residuals of the ARIMA(0, 1, 3)-model fitted to κ_t from the Poisson-gamma Lee-Carter model for females

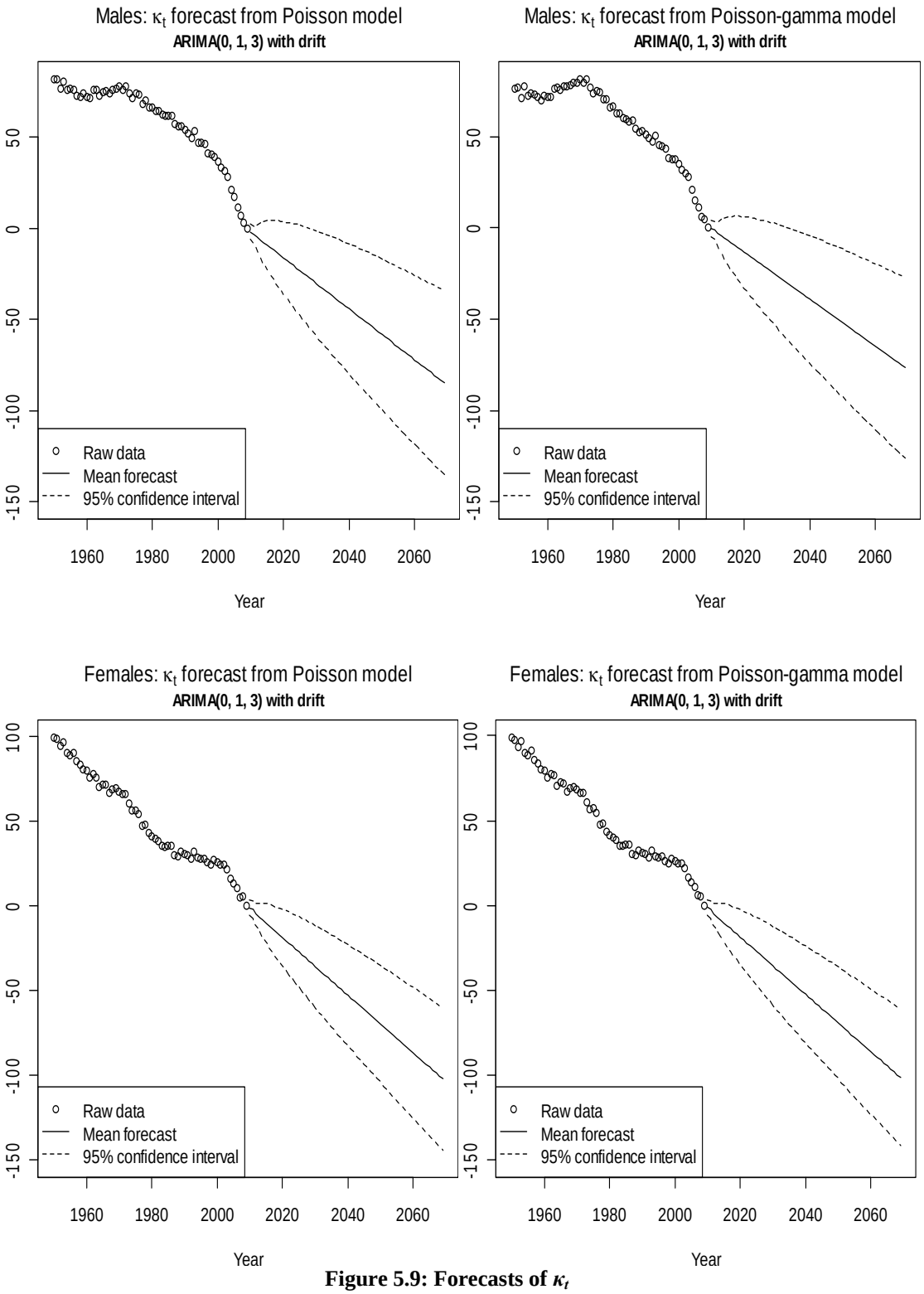

Figure 5.9: Forecasts of κ_t

Results for the male mortality index with second order integrated ARIMA-model

As mentioned earlier the κ_t's for males show a slight downward quadratic trend indicating that a second order integrated ARIMA-model would give a better fit. Using the same approach as above we estimate second order integrated ARIMA-models for the Poisson and Poisson-gamma model. The estimation results are shown in Table 5.7, Figure 5.10 and Figure 5.11, while Figure 5.12 depicts the corresponding forecasts. We observe that not only the mean forecast of κ_t shows a stronger decrease, but also that the 95% interval widens considerably compared to the forecasts obtained from first order integrated ARIMA-models. Hence, using second order integrated ARIMA-models would imply that the male mortality rate could diverge from the female mortality rate. We do not find such an implication biologically sound and therefore we discard second order integrated ARIMA-models for males.

	Poisson	**Poisson-gamma**
ARIMA	**(3, 2, 3)**	**(0, 2, 2)**
$\boldsymbol{\varphi_1}$	-0.9464	
(s.e.)	*(0.1936)*	
$\boldsymbol{\varphi_2}$	-1.1562	
(s.e.)	*(0.1394)*	
$\boldsymbol{\varphi_3}$	-0.3760	
(s.e.)	*(0.1658)*	
$\boldsymbol{\theta_1}$	-0.3116	-1.2374
(s.e.)	*(0.1669)*	*(0.1241)*
$\boldsymbol{\theta_2}$	0.6316	0.4675
(s.e.)	*(0.1387)*	*(0.1156)*
$\boldsymbol{\theta_3}$	-0.6033	
(s.e.)	*(0.1513)*	
AIC	260.55	269.04
BIC	274.97	275.22

Table 5.7: Second order integrated ARIMA-models fitted to κ_t for males

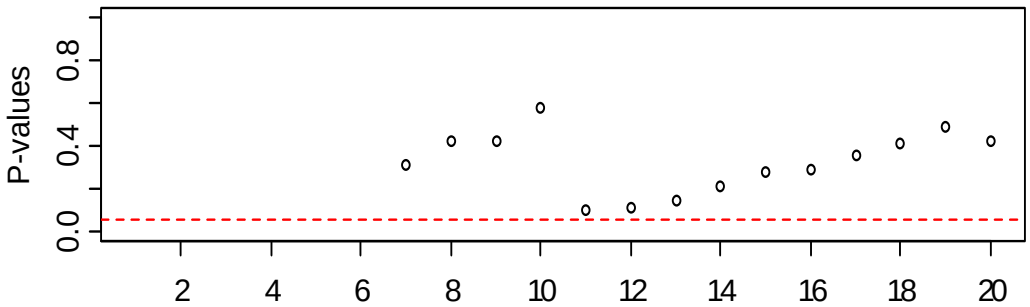

Figure 5.10: Ljung-Box *Q*-test performed on the residuals of the ARIMA(3, 2, 3)-model fitted to κ_t from the Poisson Lee-Carter model for males

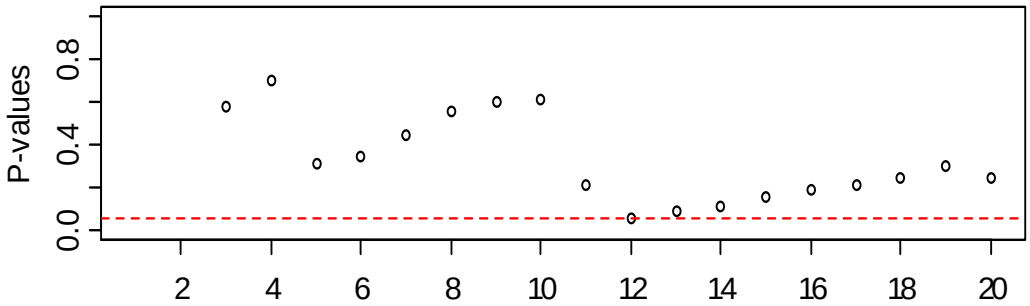

Figure 5.11: Ljung-Box *Q*-test performed on the residuals of the ARIMA(0, 2, 2)-model fitted to κ_t from the Poisson- gamma Lee-Carter model for males

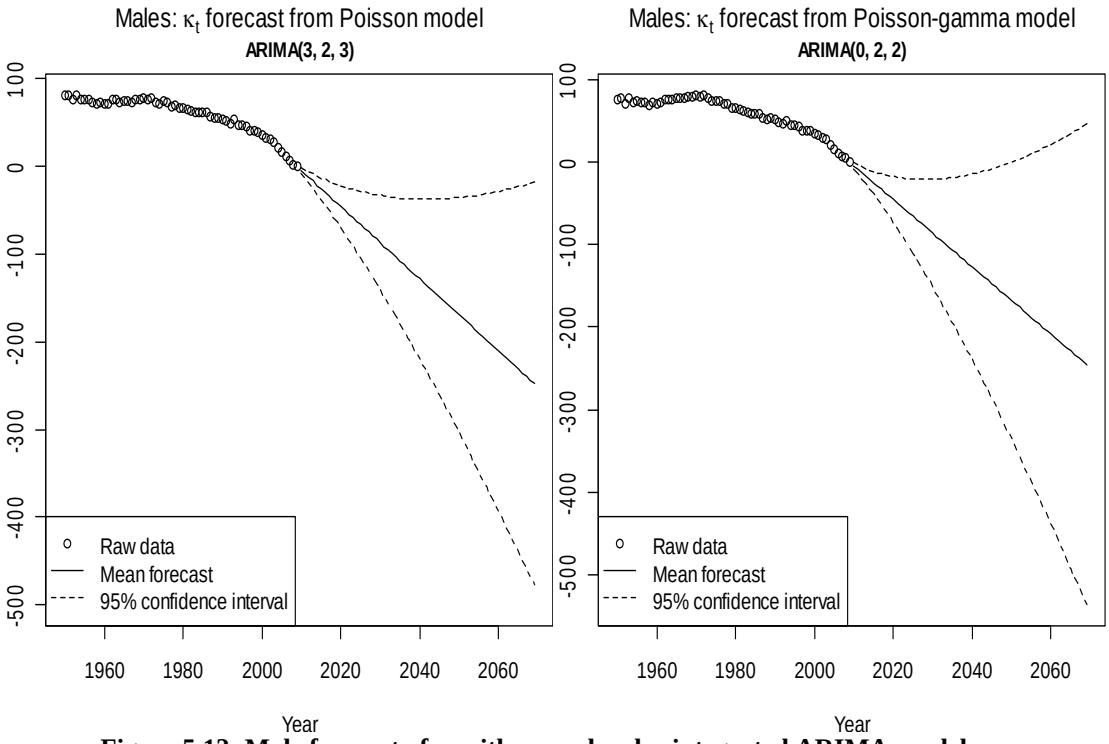

Figure 5.12: Male forecast of κ_t with second order integrated ARIMA-models

5.5 Results of the Dutch mortality forecast

In this section we forecast the future Dutch mortality rate for the next 50 years, i.e. for the years 2010 to 2059. To this end we perform simulations with the Poisson and Poisson-gamma Lee-Carter model as described in section 2.5 respectively section 3.5. The ARIMA-models estimated in section 5.4 are used for the simulations. Note that the results have been obtained by applying the jump-off correction.

Figure 5.13 depicts the forecasted age profile at year 2059 for the male respectively female Dutch population together with their 95% confidence intervals. We observe that the mean forecast of the Poisson-gamma model lies close to the mean forecast of the Poisson model, which we consider desirable.

Figure 5.14 shows the results of the mortality forecast for males together with their 95% confidence intervals. Forecasts are given for ages $x = 25, 45, 65$ and 90. The

corresponding results for females are shown in Figure 5.15. Concerning the male mortality, the Poisson-gamma model results in a higher mean forecast than the Poisson model, while for the female mortality the mean forecasts of both models almost coincide. Next, we observe that for both models the mean forecast of female mortality rate decreases faster than for males. These results correspond with the forecast results of κ_t obtained from the ARIMA-models described in section 5.4. The forecast of κ_t gives an indication for the forecast of the mortality rate, because in the Lee-Carter model the mortality index κ_t acts as the only driver for the mortality development.

To compare the interval forecasts of the Poisson-gamma model with the Poisson model we calculate the relative increase in width of the interval forecasts averaged over the projection years. We first calculate the absolute width of the interval forecast for age x in projection year t as follows

$$W_{x,t} = m_{x,t}^{U} - m_{x,t}^{L} \tag{5.6}$$

where $m_{x,t}^{U}$ and $m_{x,t}^{L}$ denote the upper respectively lower bound of the interval forecast. Then the average relative increase over the projection years is calculated as

$$\text{Average relative increase for age } x = \frac{\sum_{t=t_1}^{t_n} \left(\frac{W_{x,t}^{PG}}{W_{x,t}^{P}} - 1 \right)}{t_n - t_1 + 1} \tag{5.7}$$

where P and PG refer to the Poisson respectively Poisson-gamma model, and t_1 and t_n denote the first respectively last projection year. The resulting values are listed in the next table.

	Average relative increase in width of the interval forecasts	
	Males	**Females**
Age 25	5.38%	-0.15%
Age 45	1.26%	1.63%
Age 65	19.52%	3.86%
Age 90	44.93%	9.20%

Table 5.8: Average relative increase in width of the interval forecasts of the Poisson-gamma model compared to the Poisson model

For the male population the interval forecasts obtained by the Poisson-gamma model are significantly wider than obtained by the Poisson model for the ages 25, 65 and 90. For females the interval forecasts given by the Poisson-gamma model are wider than the Poisson model for the ages 45, 65 and 90. For age 25 the interval forecast surprisingly becomes narrower, but the decrease in width does not seem significant. These results support that the Poisson-gamma model is able to capture more variability by relaxing the mean-variance equality restriction which is imposed in the Poisson model. From the results we also conclude that the Dutch male mortality data contain more overdispersion than for females.

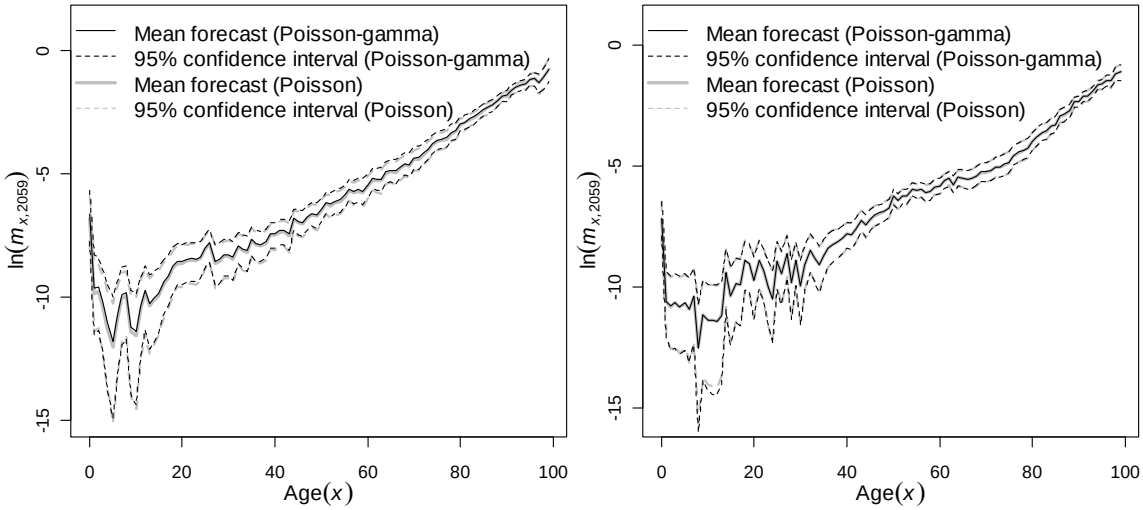

Figure 5.13: Forecasted age profile at year 2059 for males (left panel) and females (right panel)

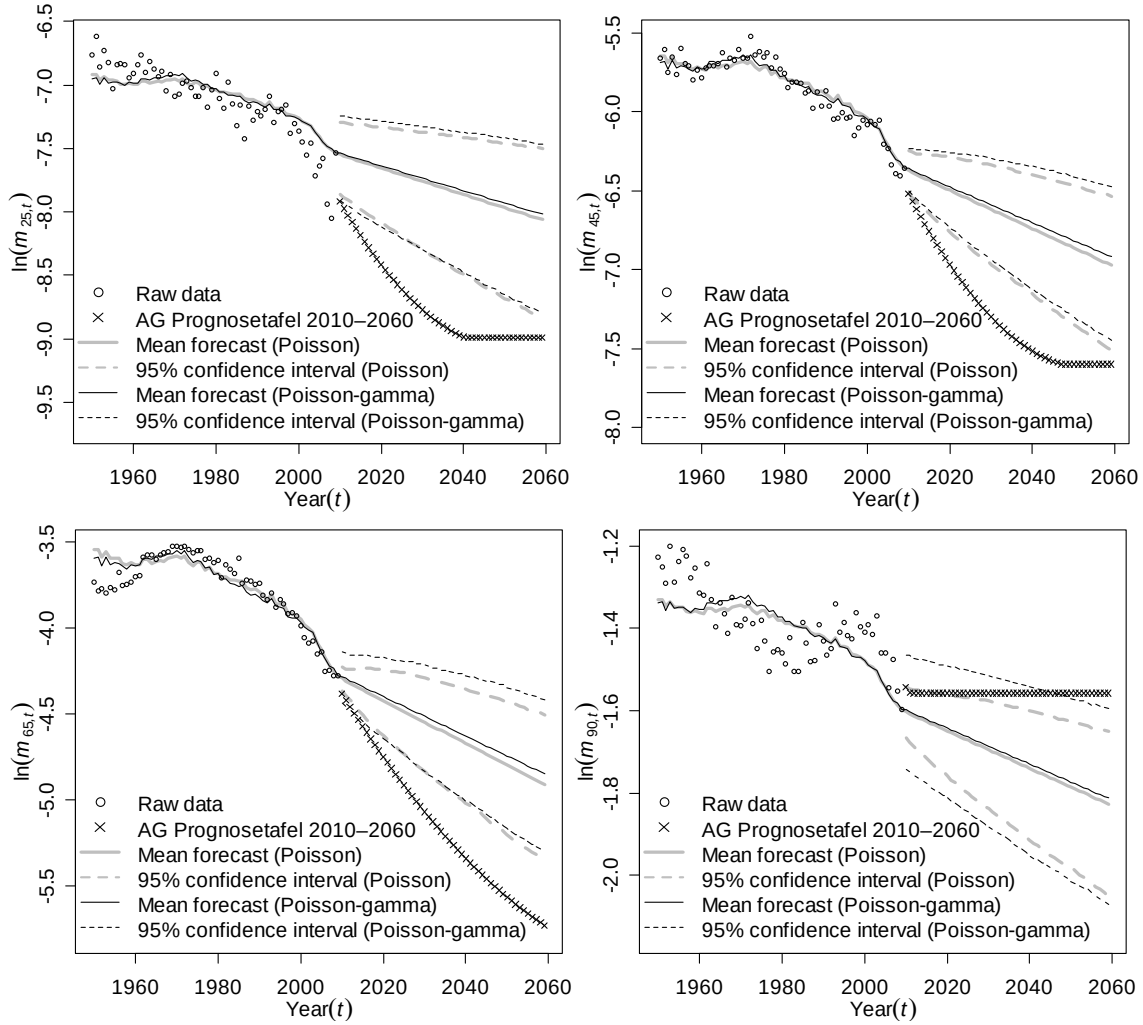

Figure 5.14: Mortality forecast for males

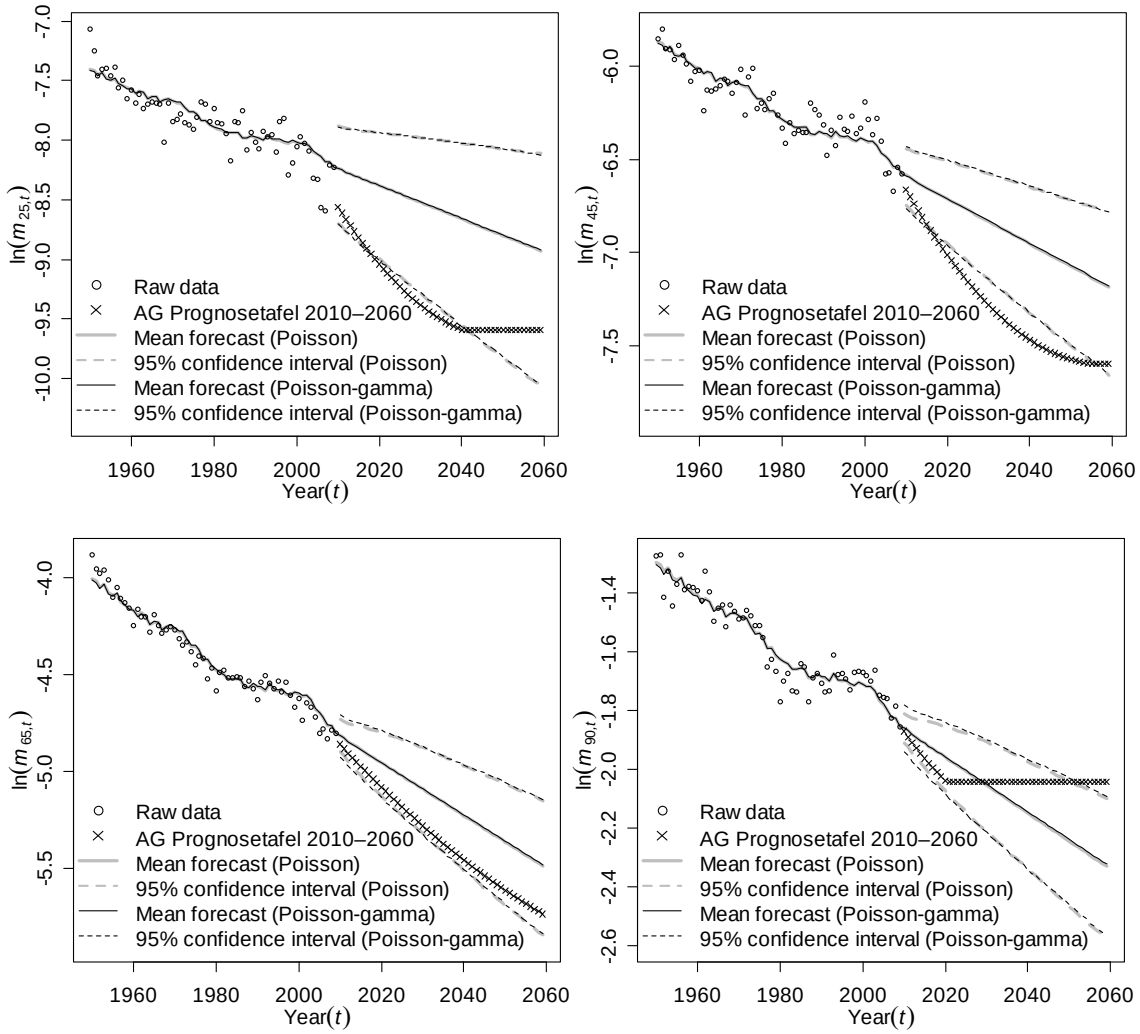

Figure 5.15: Mortality forecast for females

5.6 Backtesting

To evaluate the Poisson and Poisson-gamma Lee-Carter models we use the method of backtesting, which assesses the forecast quality of a model as if it actually had been applied in the past. In our approach we estimate both models using the historical mortality data till the year 1999 which is referred to as the training set. The data set containing information from the remaining years 2000 to 2009 is referred to as the validation set. The evaluation of the forecast quality consists of assessing whether the interval forecasts include the mortality rates of the validation set.

Figure 5.16 depicts the backtesting results for the male population for the ages $x = 25, 45, 65$ and 90. For both models we observe that the interval forecasts do not include all the mortality rates of the validation set, especially in the later years. The cause of this weak performance lies in the huge mortality improvements experienced by the male population in the recent years. As a consequence the κ_t's show a slight downward quadratic trend as mentioned in section 5.4. Even though a second order integrated ARIMA-model would give a better fit, we explicitly chose for a first order integrated ARIMA-model by reason of biological soundness. We conclude that this choice results in

a model unable to forecast the recent fast improvements in male mortality rates in both the Poisson and the Poisson-gamma setting. Additionally we would like to note that for age 90 the fitted and forecasted mortality are remarkably constant which is not in line with the results of the other ages and the results of the previous section depicted in Figure 5.14. The cause of this strange behavior lies in the jump-off correction restricting the fit and forecast to go through the last observation, which in this particular case leads to unusual results.

Figure 5.17 shows the results of the backtesting for the female mortality for the ages $x = 25, 45, 65 \text{ and } 90$. For age 45 the interval forecasts of both models contain the mortality rates of the complete validation set, while this is not the case for the other ages. For age 25 the backtesting results show no significant difference between both models since their interval forecasts almost coincide. For age 65 and 90 several mortality rates of the validation set fall outside the interval forecast given by the Poisson model, but are still contained by the Poisson-gamma model due to wider interval forecasts. In conclusion, both models have difficulties forecasting the future mortality rate for females, and on the whole the Poisson-gamma model has higher forecast quality than the Poisson model.

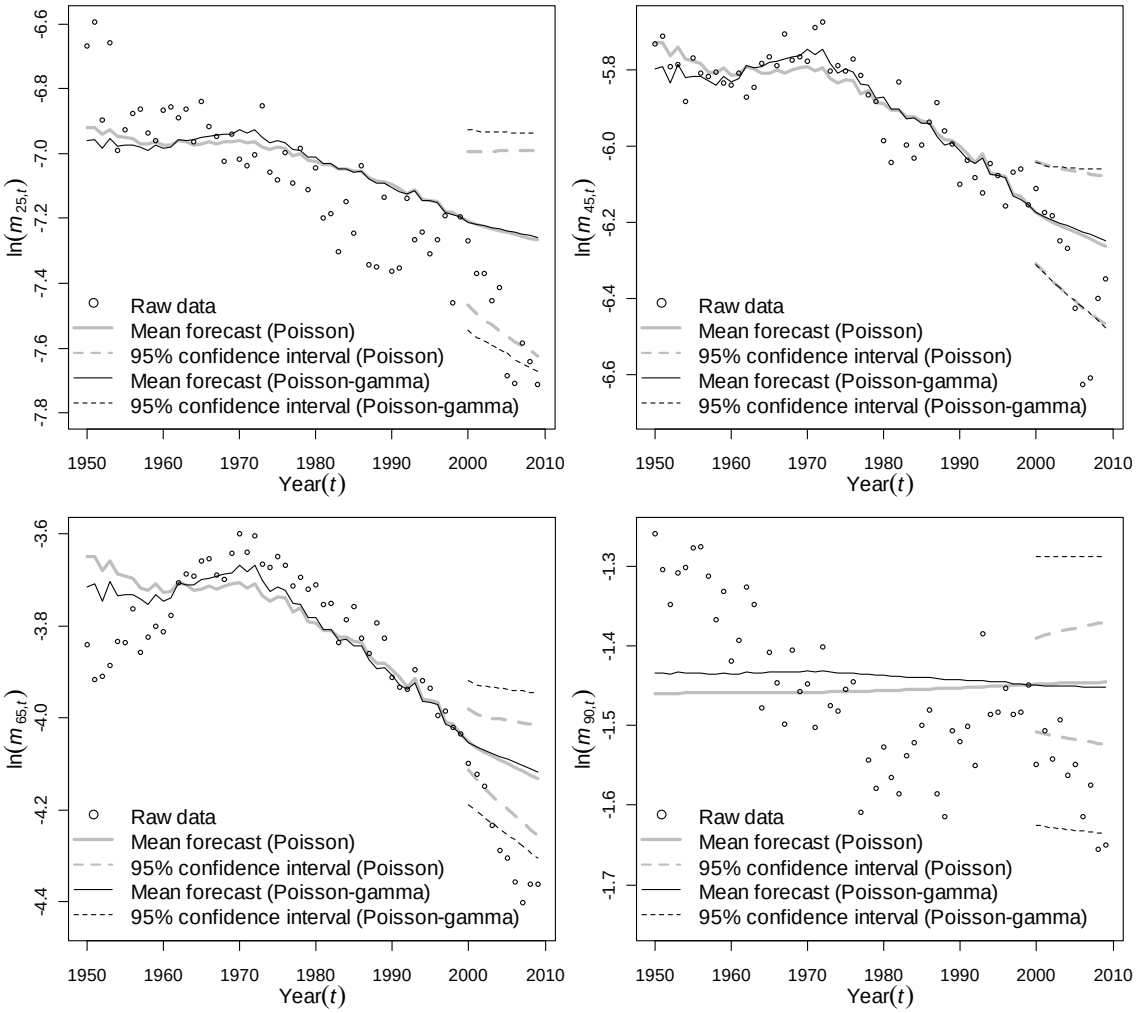

Figure 5.16: Backtesting results for males

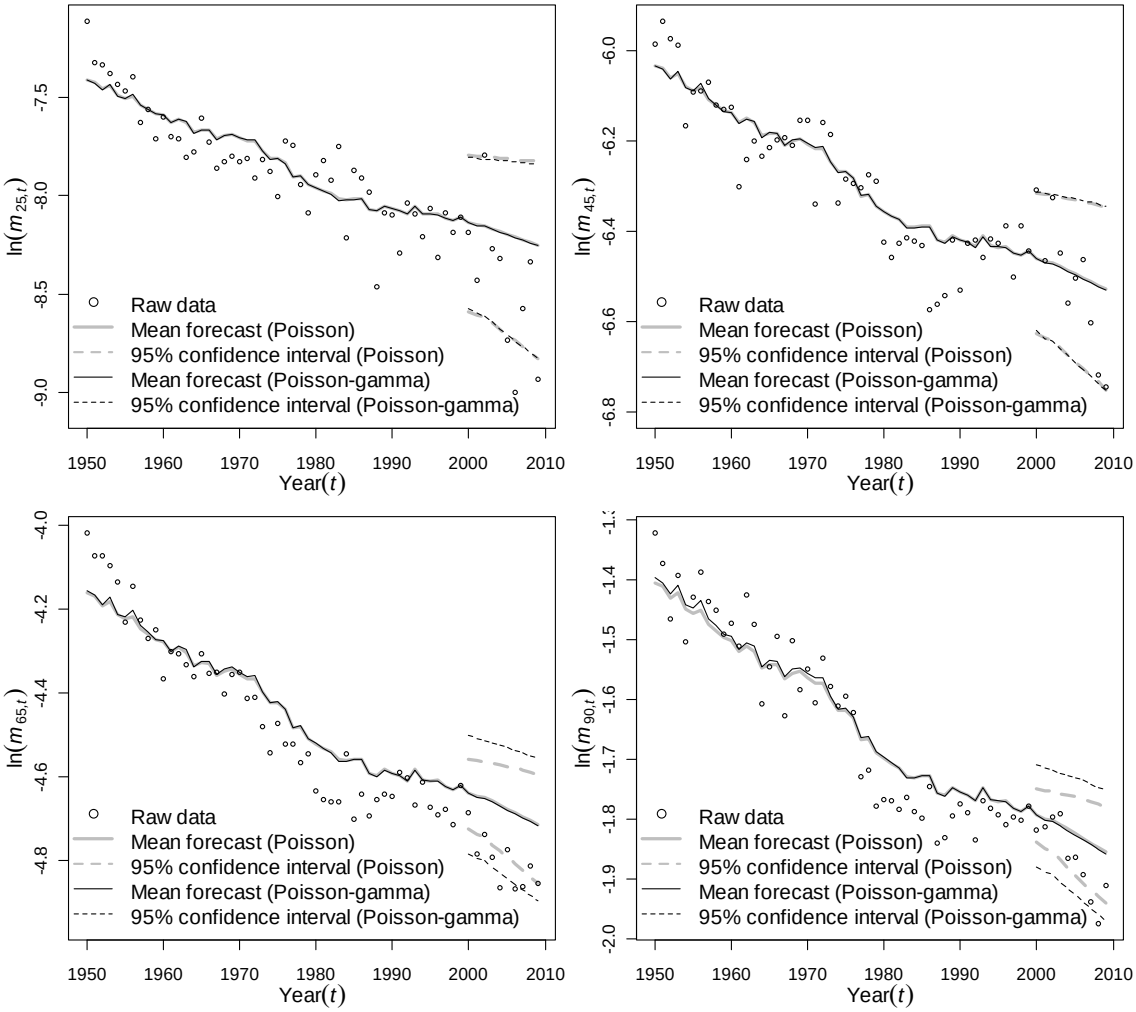

Figure 5.17: Backtesting results for females

5.7 Comparison with "AG Prognosetafel 2010–2060"

In 2010 the Dutch Actuarial Association (Actuarieel Genootschap) published a mortality table containing a mortality forecast for 2010 to 2060 generally known as the "AG Prognosetafel 2010-2060" (*Prognosetafel 2010-2060*, 2010). For convenience, we refer to this mortality table as the "AG-table". The Dutch Actuarial Association created the AG-table using the raw mortality data collected by Statistics Netherlands (Centraal Bureau voor de Statistiek). The model used by the Dutch Actuarial Association distinguishes for the mortality development a short-term trend and a long-term trend. To represent the long-term trend the model uses a so-called goal mortality table, which is the forecasted mortality table for the year 2060 based on the historical data from the years 1988 to 2008. To do justice to the recent mortality development the model forecasts a short-term trend based on the observed mortality from 2001 to 2008. Note that the short-term trend does not affect the forecasted goal mortality table for 2060, but only the speed of convergence towards this table.

The AG-table contains the single-year death probability $q_{x,t}$, which is the probability that a person aged x last birthday in year t will die within one year. The Lee-Carter model

is based on the mortality rate $m_{x,t}$, which is the instantaneous death rate for a person aged x at time t (also known as the force of mortality). The relationship between the single-year death probability $q_{x,t}$ and the mortality rate $m_{x,t}$, is given by

$$q_{x,t} = 1 - e^{-\int_0^1 m_{x+s,t+s} ds} \tag{5.8}$$

For the derivation we refer the interested reader to Bowers *et al.* (1997). Additionally, we assume that for any integer age x and calendar year t the following holds.

$$m_{x+\xi,t+\tau} = m_{x,t} \quad \text{for } 0 \leq \xi, \tau < 0 \tag{5.9}$$

The assumption above basically states that the mortality rate remains constant within an age-period cell. Using this assumption equation (5.8) simplifies to

$$q_{x,t} = 1 - e^{-m_{x,t}} \tag{5.10}$$

To compare the forecasts of the AG-table with the forecasts obtained by the Poisson and Poisson-gamma Lee-Carter model, we transform the death probabilities of the AG-table to mortality rates using equation (5.10).

Figure 5.14 shows the three forecasts for males for age $x = 25, 45, 65$ and 90. For age 25, 45, and 65 the short-term as well as the long-term trend of the AG-table is not included by the interval forecasts. For age 90 the AG-table shows a flat mortality development, which for a large part is included by the interval forecasts of the Poisson-gamma model, but not by the Poisson model.

Figure 5.15 contains the comparison of the three mortality forecasts for females for age $x = 25, 45, 65$ and 90. For the ages 25 and 45 the short-term trend of the AG-table falls outside the interval forecasts, but eventually the long-term trend is included in the interval forecasts. For age 65 both trends are included by the interval forecasts. For age 90 the mortality development as stated by the AG-table falls outside the interval forecasts for the first couple of years and then is included in the interval forecasts, but eventually the mortality rate falls outside the interval forecasts again.

We conclude for males that the difference in modeling, together with the fast mortality improvements in recent years, leads to significant differences between the AG-table and the forecasts of the Poisson and Poisson-gamma model for both the short-term and the long-term trend (except for age 90). The female population also experienced relatively fast mortality improvements in the recent years, but not as fast as the male population. Therefore the AG-table differs from the forecasts of the Poisson and Poisson-gamma model in the short-term, but this difference becomes smaller in the long-term and often diminishes.

Chapter 6

Bayesian modeling of the portfolio mortality using the Poisson-gamma setting

Determining the mortality rate of a portfolio of insured lives often proves to be difficult. The amount of historical portfolio data is limited in terms of the size of the dataset as well as the number of years of portfolio data. Data scarcity forms the main obstacle to fit a stochastic mortality model reliably. Usually sufficient historical data is available to determine the mortality rate of a country population. A commonly applied approach within the actuarial field to overcome the obstacle of limited data consists in estimating the portfolio mortality through an experience factor. The portfolio experience factor reflects how much the country population mortality rate needs to be adjusted for the portfolio mortality rate.

Essentially, the experience factor method incorporates external information which subsequently is adjusted according to one's historical portfolio data. In the insurance industry several ad-hoc techniques to calculate the experience factor are applied, see *Generatietafels Pensioenen 2010* (2010). Bayesian probability theory provides a tool to achieve this in a formal mathematical way. Olivieri & Pitacco (2009) proposed an approach to determine a portfolio experience factor with Bayesian inference techniques using the conjugate property of the Poisson-gamma distribution. We recognize that this method can be applied to the Poisson-gamma Lee-Carter model in a natural way allowing us to use the full forecasting ability of the Lee-Carter model for portfolio mortality.

Section 6.1 gives an introduction to Bayesian probability theory. Section 6.2 discusses the conjugate property of the Poisson-gamma distribution. Section 6.3 presents the Bayesian estimation method for the portfolio experience factor which utilizes the conjugate property within the Poisson-gamma Lee-Carter model. Section 6.4 derives the mathematical properties of the Bayesian estimator. Section 6.5 discusses the forecasting of mortality for the Bayesian extension and section 6.6 presents the simulation method to obtain the corresponding confidence intervals.

6.1 Introducing Bayesian probability theory

The main difference between classical statistics and Bayesian statistics lies in the treatment of the parameters of a statistical model. While classical statistics treats the parameters as having fixed but unknown values, Bayesian statistics considers the parameters of a statistical model to be random variables, thus having a probability distribution. This allows reasoning about the probability of the parameters.

Equivalent to classical statistics, one first starts by specifying a statistical model for the observed data $Y = y$ given an unknown vector of parameters θ.

$$f(Y|\theta) \tag{6.1}$$

In addition to the probability distribution above, Bayesian statistics considers the parameter θ to be a random variable having distribution $\pi(\theta|\eta)$. This distribution is referred to as the prior distribution of θ, because it expresses the probability distribution assigned to θ prior to having observed any data. The probability distribution of θ possesses its own vector of parameters η, which are referred to as hyperparameters. We consider η to be a constant and therefore suppress it in the notations. Finally, conditional on the observed data inferences about the probability distribution of θ can be made by applying Bayes' Theorem as follows

$$\pi(\theta|y) = \frac{f(y,\theta)}{f(y)} = \frac{f(y,\theta)}{\int f(y,\theta')d\theta'} = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta')\pi(\theta')d\theta'} \tag{6.2}$$

The distribution above is called the posterior distribution, because it expresses the probability distribution assigned to θ after having observed the data. Note that the denominator of the posterior distribution acts as a normalizing constant. Hence, the posterior distribution is proportional to the likelihood times the prior.

$$\pi(\theta|y) \propto f(y|\theta)\pi(\theta) \tag{6.3}$$

In conclusion, Bayesian statistics allows one to formulate a prior distribution which reflects one's beliefs towards the parameter θ before having seen any data. These prior beliefs can come from a different, possibly external source, e.g. expert judgment. After observing the data one obtains the posterior distribution, which essentially reflects the updated beliefs towards the parameter θ based on a combination of the prior beliefs and the observed data. For a more in-depth theoretical treatment of Bayesian theory we refer the interested reader to Ghosh *et al.* (2006) or Carlin & Louis (2000).

6.2 The conjugate property of the Poisson-gamma distribution

Calculating the integral in the denominator of the posterior distribution (6.2) has been the main obstacle to the practical use of Bayesian statistics. Recently this obstacle has been overcome due to the introduction of the Markov chain Monte Carlo (MCMC) methodology (see Gilks *et al.* (1996)) and the arrival of generic software for MCMC, such as WinBUGS and JAGS. In our thesis we do not resort to MCMC-methods, but instead we rely on the conjugate property of the Poisson-gamma distribution.

Choosing the prior distribution $\pi(\theta)$ such that it is conjugate to the likelihood $f(y|\theta)$, leads to the posterior distribution $\pi(\theta|y)$ coming from the same distributional family as

the prior distribution. The main advantage of this approach is that we can avoid the calculation of the integral in the denominator of the posterior distribution. Likelihood functions belonging to the exponential families do in fact have conjugate priors, which can be proven as follows. Let y_1, …, y_n be n observations from an exponential family distribution.

$$f(y_i|\theta) = s(y_i)t(\theta)e^{b(\theta)^T a(y_i)} \propto t(\theta)e^{b(\theta)^T a(y_i)} \tag{6.4}$$

where the superscript T refers to the transpose of the indicated vector. Then the corresponding likelihood is proportional as follows

$$f(\boldsymbol{y}|\theta) \propto \big(t(\theta)\big)^n e^{b(\theta)^T \sum_i a(y_i)} \tag{6.5}$$

If we construct a conjugate family of prior distributions with hyperparameters η_1 and η_2 in the following way

$$\pi(\theta|\eta_1, \eta_2) \propto \big(t(\theta)\big)^{\eta_1} e^{b(\theta)^T \eta_2} \tag{6.6}$$

then the posterior distribution is

$$\pi(\theta|\boldsymbol{y}) \propto f(\boldsymbol{y}|\theta)\pi(\theta|\eta) = \big(t(\theta)\big)^{n+\eta_1} e^{b(\theta)^T (\sum_i a(y_i)+\eta_2)} \tag{6.7}$$

Except from a normalizing constant denominator, we recognize the latter to be a member of the exponential family.

In our thesis we will specifically focus on the likelihood from the Poisson distribution and its conjugate prior, the gamma distribution. For observed data $Y = y$ the Poisson likelihood is given by

$$f(y|\theta) = e^{-\theta}\frac{\theta^y}{y!} \tag{6.8}$$

and the conjugate prior coming from the gamma distribution $Gamma(\alpha, \beta)$ is given by

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta} \tag{6.9}$$

The posterior distribution can be determined to be proportional to

$$\pi(\theta|y) \propto f(y|\theta)\pi(\theta) = e^{-\theta}\frac{\theta^y}{y!}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta} \propto \theta^{(\alpha+y)-1}e^{-(\beta+1)\theta} \tag{6.10}$$

Except from a normalizing constant denominator, we recognize that the posterior distribution comes from the gamma distribution $Gamma(\alpha + y, \beta + 1)$. Hence, after having observed y the parameters of the prior gamma distribution are updated to $\alpha' = \alpha + y$ and $\beta' = \beta + 1$.

6.3 Bayesian estimation of the portfolio experience factor

This section discusses a method to determine the portfolio experience factor with Bayesian inference techniques. First we assume that the Poisson-gamma Lee-Carter model as specified in chapter 3 is fitted to the country population mortality data, which gives us the parameter estimates of $\{\alpha_x\}$, $\{\beta_x\}$, $\{\kappa_t\}$ and $\{\bar{\varphi}_x\}$. Thus the model for the country population mortality is given by

$$\begin{cases} D_{x,t}^{pop}|\bar{Z}_x^{pop} \sim Poisson\big(e_{x,t}^{pop} m_{x,t}^{pop} \bar{Z}_x^{pop}\big) \\ \bar{Z}_x^{pop} \sim Gamma(\bar{\varphi}_x^{-1}, \bar{\varphi}_x^{-1}) \end{cases} \tag{6.11}$$

where $m_{x,t}^{pop} = e^{\alpha_x + \beta_x \kappa_t}$. We use the superscript *pop* and *port* to distinguish between population and portfolio mortality. In the mortality model above the random variable $\bar{Z}_x^{pop}$ acts as an age-specific experience factor with the imposed assumption $E\left[\bar{Z}_x^{pop}\right] = 1$. This assumption ensures that $\bar{Z}_x^{pop}$ does not affect the mean mortality rate, which is desirable in case of country population mortality.

Next, we specify the Poisson-gamma Lee-Carter model for portfolio mortality as follows

$$\begin{cases} D_{x,t}^{port} | \bar{Z}_x^{port} \sim Poisson\left(e_{x,t}^{port} m_{x,t}^{pop} \bar{Z}_x^{port}\right) \\ \bar{Z}_x^{port} \sim Gamma(\bar{\varphi}_x^{-1}, \bar{\varphi}_x^{-1}) \end{cases} \tag{6.12}$$

where we treat $m_{x,t}^{pop}$ as a constant, estimated by model (6.11). Analogously $\bar{Z}_x^{port}$ acts as an age-specific experience factor, but in this case for the portfolio mortality rate. In the context of Bayesian probability theory one can also view $\bar{Z}_x^{port}$ as a stochastic parameter with prior distribution having the same hyperparameters as estimated by model (6.11) for the country population mortality. In other words, our prior beliefs concerning the experience factor for the portfolio stem from the model fitted for the country population mortality.

To determine the posterior distribution of the age-specific experience factor $\bar{Z}_x^{port}$ we rely on the conjugate property of the Poisson-gamma distribution as described in section 6.2. The stochastic parameter $\left(e_{x,t}^{port} m_{x,t}^{pop} \bar{Z}_x^{port}\right)$ of the Poisson distribution has the following prior distribution

$$e_{x,t}^{port} m_{x,t}^{pop} \bar{Z}_x^{port} \sim Gamma\left(\bar{\varphi}_x^{-1}, \left(e_{x,t}^{port} m_{x,t}^{pop} \bar{\varphi}_x\right)^{-1}\right) \tag{6.13}$$

Let us denote with $d_{x,t}^{port}$ the observed numbers of death for age x in year t. After having observed $d_{x,t}^{port}$ we determine the posterior distribution to be

$$e_{x,t}^{port} m_{x,t}^{pop} \bar{Z}_x^{port} | d_{x,t}^{port} \sim Gamma\left(\bar{\varphi}_x^{-1} + d_{x,t}^{port}, \left(e_{x,t}^{port} m_{x,t}^{pop} \bar{\varphi}_x\right)^{-1} + 1\right) \tag{6.14}$$

Finally, the posterior distribution of $\bar{Z}_x^{port}$ is given by

$$\bar{Z}_x^{port} | d_{x,t}^{port} \sim Gamma\left(\bar{\varphi}_x^{-1} + d_{x,t}^{port}, \bar{\varphi}_x^{-1} + e_{x,t}^{port} m_{x,t}^{pop}\right) \tag{6.15}$$

The posterior distribution of $\bar{Z}_x^{port}$ above can be generalized for more observations. Let us denote the observations of the death count in the historical portfolio data as $d_{x,t}^{port}$ for age x for the years $t = 0, \dots, T$. The posterior distribution of $\bar{Z}_x^{port}$ can then be generalized as follows

$$\bar{Z}_x^{port} \left| d_{x,0}^{port}, \dots, d_{x,T}^{port} \sim Gamma\left(\bar{\varphi}_x^{-1} + \sum_{t=0}^{T} d_{x,t}^{port}, \bar{\varphi}_x^{-1} + \sum_{t=0}^{T} e_{x,t}^{port} m_{x,t}^{pop}\right)\right. \tag{6.16}$$

Using the results derived in section 3.1 we conclude that the unconditional distribution of the portfolio death count after updating corresponds with

$$D_{x,t}^{port} \left| d_{x,0}^{port}, \dots, d_{x,T}^{port} \sim NegBin\left(\bar{\varphi}_x^{-1} + \sum_{t^*=0}^{T} d_{x,t^*}^{port}, \frac{\bar{\varphi}_x^{-1} + \sum_{t^*=0}^{T} e_{x,t^*}^{port} m_{x,t^*}^{pop}}{\bar{\varphi}_x^{-1} + \sum_{t^*=0}^{T} e_{x,t^*}^{port} m_{x,t^*}^{pop} + e_{x,t}^{port} m_{x,t}^{pop}}\right)\right. \tag{6.17}$$

In conclusion, by fitting the Poisson-gamma Lee-Carter model to the to the country population mortality data we do not only obtain the estimate for the country population mortality rate $m_{x,t}^{pop}$, but we also obtain the parameter estimate of $\bar{\varphi}_x$. From a Bayesian perspective the latter parameter estimate can be viewed as the hyperparameter belonging to the prior distribution of the portfolio experience factor $\bar{Z}_x^{port}$, i.e. the prior beliefs stem from the country population mortality. The posterior distribution of $\bar{Z}_x^{port}$ can then be determined with Bayesian inference techniques.

The Bayesian approach as described above has several desirable features. First, we do not use MCMC-methods which makes our approach relatively simple, easy to understand and implement. Next, the approach can deal effectively with a portfolio which has limited amount of historical data or even missing data. If little data is available for a certain age *x*, the posterior distribution will be close to the prior distribution. Since the approach uses a prior distribution based on the country population mortality, the portfolio mortality rate will be close to the country population mortality rate in such circumstances, which can be considered as the best available estimate.

6.4 Properties of the Bayesian estimator of the portfolio experience factor

In this section we analyze the mean and the variance of the Bayesian estimator of the portfolio experience factor given by equation (6.16). Using the well-known properties of the gamma distribution we derive the posterior mean and posterior variance to be

$$E\left[\bar{Z}_x^{port}\middle|d_{x,0}^{port},\dots,d_{x,T}^{port}\right]=\frac{\bar{\varphi}_x^{-1}+\sum_{t=0}^{T}d_{x,t}^{port}}{\bar{\varphi}_x^{-1}+\sum_{t=0}^{T}e_{x,t}^{port}m_{x,t}^{pop}} \tag{6.18}$$

$$Var\left(\bar{Z}_x^{port}\middle|d_{x,0}^{port},\dots,d_{x,T}^{port}\right)=\frac{\bar{\varphi}_x^{-1}+\sum_{t=0}^{T}d_{x,t}^{port}}{\left(\bar{\varphi}_x^{-1}+\sum_{t=0}^{T}e_{x,t}^{port}m_{x,t}^{pop}\right)^2} \tag{6.19}$$

The posterior expectation of the portfolio experience factor is a function of the total observed number of deaths implying that the Bayesian estimator disregards the timing of the death occurrence. From equation (6.18) we derive that $E\left[\bar{Z}_x^{port}\middle|d_{x,0}^{port},\dots,d_{x,T}^{port}\right]>1$ if and only if $\sum_{t=0}^{T}d_{x,t}^{port}>\sum_{t=0}^{T}e_{x,t}^{port}m_{x,t}^{pop}$. In other words, the mean of the Bayesian estimator of the portfolio experience factor has value greater than 1 if and only if the observed number of deaths is greater than the expected number of deaths based on the prior assumptions, i.e. population mortality rate. Hence, the portfolio experience factor gets adjusted based on the discrepancy between the observations and the expectations.

The mean of the Bayesian estimator can be rewritten in the following form

$$\begin{cases} E\left[\bar{Z}_x^{port}\middle|d_{x,0}^{port},\dots,d_{x,T}^{port}\right]=fE\left[\bar{Z}_x^{port}\right]+(1-f)\dfrac{\sum_{t=0}^{T}d_{x,t}^{port}}{\sum_{t=0}^{T}e_{x,t}^{port}m_{x,t}^{pop}} \\ \text{with } f=\dfrac{\bar{\varphi}_x^{-1}}{\bar{\varphi}_x^{-1}+\sum_{t=0}^{T}e_{x,t}^{port}m_{x,t}^{pop}} \end{cases} \tag{6.20}$$

Thus the posterior mean of the portfolio experience factor is a weighted average of its prior expectation $E\left[\bar{Z}_x^{port}\right]$ and the average observed portfolio experience factor $\frac{\sum_{t=0}^{T} d_{x,t}^{port}}{\sum_{t=0}^{T} e_{x,t}^{port} m_{x,t}^{pop}}$. Such a form is well-known from credibility theory where *f* is called the credibility factor.

The posterior mean has the following asymptotic properties

$$\lim_{\bar{\varphi}_x \to \infty} E\left[\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right] = \frac{\sum_{t=0}^{T} d_{x,t}^{port}}{\sum_{t=0}^{T} e_{x,t}^{port} m_{x,t}^{pop}} \tag{6.21}$$

$$\lim_{\bar{\varphi}_x \to 0} E\left[\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right] = 1 \tag{6.22}$$

and the posterior variance

$$\lim_{\bar{\varphi}_x \to \infty} Var\left(\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right) = \frac{\sum_{t=0}^{T} d_{x,t}^{port}}{\left(\sum_{t=0}^{T} e_{x,t}^{port} m_{x,t}^{pop}\right)^2} \tag{6.23}$$

$$\lim_{\bar{\varphi}_x \to 0} Var\left(\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right) = 0 \tag{6.24}$$

Equation (6.18) and the asymptotic properties above state that the sensitivity of the portfolio experience factor to adjustments is an increasing function of φ_x. Furthermore, φ_x comes from the Poisson-gamma Lee-Carter model fitted to the country population mortality data and acts as an age-specific dispersion parameter explicitly capturing unexplained heterogeneity. Intuitively the relation between φ_x and the sensitivity can be explained as follows. The more unexplained heterogeneity the country population mortality data contain, the less reliable these data are considered. Therefore more value is attached to the portfolio observations, which results in the portfolio observations having a larger impact on the adjustment of the portfolio experience factor.

6.5 Mortality forecasting

After having obtained the posterior distribution of the portfolio experience factor $\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}$ as described in section 6.3, we can forecast the portfolio mortality relative to the country population mortality forecast. Let us denote the *s*-period ahead forecast of the country population mortality rate as $\hat{m}_{x,t_n+s}^{pop}$ which is obtained using the mortality forecasting method described in section 3.4. Then the expected value of future portfolio death count in the Poisson-gamma Lee-Carter model is given by

$$E\left[D_{x,t_n+s}^{port}\right] = e_{x,t_n+s}^{port} \hat{m}_{x,t_n+s}^{pop} E\left[\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right] \tag{6.25}$$

where e_{x,t_n+s}^{port} is the future exposure and $E\left[\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right]$ acts as the mean posterior portfolio experience factor. The latter is obtained by taking the mean of the posterior distribution of the experience factor $\bar{Z}_x^{port}$ given by equation (6.16). Hence, the portfolio mortality rate can be defined as

$$\hat{m}_{x,t_n+s}^{port} = \hat{m}_{x,t_n+s}^{pop} E\left[\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right] \tag{6.26}$$

6.6 Simulation and confidence intervals

For the same reasons as given in section 2.5 we would like to use simulation to measure the uncertainty around the quantities of interest, which are calculated using the forecasted portfolio mortality rate. A relatively simple approach would be to calculate the mean of the posterior portfolio experience factor $E\left[\bar{Z}_x^{port} \middle| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right]$ once on the basis of the original data. Subsequently, this portfolio experience factor needs to be applied to the simulated country population mortality rate, where the latter is simulated with the semi-parametric bootstrap method described in section 3.5. However, in such simulation approach the portfolio experience factor is only determined once and thus can be considered deterministic.

To take the uncertainty around the portfolio experience factor into account we propose to apply the Bayesian bootstrap method introduced by Rubin (1981). Extending the bootstrap approach described in section 3.5 with the Bayesian bootstrap for portfolio experience factor entails simulating from the posterior distribution of the portfolio experience factor (6.16).

Simulation algorithm

The extended bootstrap approach consists of the following steps:

1. Generate *N* bootstrap samples for $n = 1, \dots, N$ for:
   (a) the country population death count $d_{x,t}^{pop(n)}$, which are realizations from the negative binomial distribution given by (3.12);
   (b) the portfolio experience factor $\bar{z}_x{}^{(n)}$, which are realizations from the posterior distribution given by (6.16).
2. For the *n*th bootstrap sample:
   (a) Re-estimate the Poisson-gamma Lee-Carter model obtaining the parameter sets $\left\{\alpha_x^{(n)}\right\}, \left\{\beta_x^{(n)}\right\}, \left\{\kappa_t^{(n)}\right\}, \left\{\bar{\varphi}_x^{(n)}\right\}$;
   (b) Using $\left\{\kappa_t^{(n)}\right\}$ re-estimate the parameters of the originally estimated ARIMA model where the order of integration *d*, the degree *p* of the autoregressive polynomial and the degree *q* of the moving average polynomial do not change;
   (c) Generate a projection of future mortality index κ_t using the ARIMA model obtained in step (2b);
   (d) Calculate the forecasted future portfolio mortality rate as $\hat{m}_{x,t}^{port(n)} = \hat{m}_{x,t}^{pop(n)} \bar{z}_x{}^{(n)}$;
   (e) Calculate the quantities of interest using the forecasted future portfolio mortality rate.
3. Derive the confidence interval for the quantity of interest using the empirical distribution obtained in step (2).

Chapter 7

An application to AEGON's portfolio

In the Netherlands AEGON is one of the largest insurance companies providing a range of financial products for life insurance, pension and asset management. In this chapter we apply the Bayesian extension of the Lee-Carter model for portfolio mortality, which is presented in chapter 6, to the Dutch life and pension portfolio of AEGON.

Section 7.1 gives a description of the available historical portfolio data. Section 7.2 presents the results of the Bayesian estimation for the portfolio of AEGON. In section 7.3 we perform a sensitivity analysis for the Bayesian estimated experience factor. Section 7.4 presents the mortality forecast for the portfolio of AEGON as well as a comparison with the Dutch population mortality forecast. Due to confidentiality reasons, we omit the units of the axis in the graphs.

7.1 Description of the data

The portfolio dataset used in this chapter comes from the Dutch life and pension portfolio of AEGON. The portfolio dataset consists of historical mortality data for the years 2003 to 2009. For each year the portfolio dataset includes the following for males and females separately:

- Exposure $e_{x,t}^{port}$, measured as the number of insured aged x at the last birthday and alive at mid-calendar year t;
- Number of deaths $d_{x,t}^{port}$, measured as the number of insured aged x at the last birthday and having died in calendar year t.

From the backtesting performed in chapter 4 we concluded that the Poisson-gamma Lee-Carter model provides relatively better results for females than for males. Furthermore the proposed model for portfolio mortality can be applied on each age separately independent from other ages. In this chapter we therefore limit our focus to females for the ages 25, 45, 65 and 90.

7.2 Bayesian estimates of the portfolio experience factor

From the Poisson-Gamma Lee-Carter model fitted to the country population mortality data we obtain the prior distribution of the portfolio experience factor

$$\bar{Z}_x^{port} \sim Gamma(\bar{\varphi}_x^{-1}, \bar{\varphi}_x^{-1}) \tag{7.1}$$

Note that the prior distribution of $\bar{Z}_x^{port}$ is fitted such that $E\left[\bar{Z}_x^{port}\right] = 1$. The Bayesian estimates of the portfolio experience factor, i.e. the parameters of the posterior distribution, are obtained by updating the prior distribution with the portfolio mortality data as follows

$$\begin{cases} \bar{Z}_x^{port} \left| d_{x,2003}^{port}, \dots, d_{x,2009}^{port} \sim Gamma\left(\alpha_x^{port}, \beta_x^{port}\right) \right. \\ \alpha_x^{port} = \bar{\varphi}_x^{-1} + \sum_{t=2003}^{2009} d_{x,t}^{port} \\ \beta_x^{port} = \bar{\varphi}_x^{-1} + \sum_{t=2003}^{2009} e_{x,t}^{port} m_{x,t}^{pop} \end{cases} \tag{7.2}$$

where $e_{x,t}^{port}$ and $d_{x,t}^{port}$ denote respectively the exposure and numbers of death for age x observed in year t of the portfolio.

Table 7.1 contains the results for females of age 25, 45, 65 and 90. We observe that the posterior mean of the experience factor lies close to 1 for all ages. We conclude that the updating has little impact on the experience factor. The cause of the insensitivity to adjustments can be found in the high values of $\bar{\varphi}_x^{-1}$ as described in section 6.4.

In 2010 the Dutch Association of Insurers (Verbond van Verzekeraars) published experience factors specific for the pension insured population (*Generatietafels Pensioenen 2010*, 2010). The experience factors are derived from data provided by the majority of Dutch insurance companies and are referred to as the ES-P2 factors. For comparison Table 7.1 contains the ES-P2 factors. We observe that the obtained experience factors differ greatly from ES-P2 factors except for age 90.

Age (x)	**25**	**45**	**65**	**90**
$\bar{\varphi}_x^{-1}$	1,705,141	974.875	793.8901	931.9494
$E\left[\bar{Z}_x^{port} \left| d_{x,0}^{port}, \dots, d_{x,T}^{port}\right.\right]$	0.9999988	0.9859078	1.040745	0.9617708
ES-P2 factors	0.5835	0.7611	0.9034	0.9394

Table 7.1: Bayesian parameter estimates of the portfolio experience factor and the ES-P2 factors for females

As discussed in section 6.4, the portfolio experience factor gets adjusted based on the discrepancy between the observed number of deaths in the portfolio and the expected number of deaths. Figure 7.1 depicts these numbers for females for the ages of interest where the expected number of deaths for age x in year t is calculated as $e_{x,t}^{port} m_{x,t}^{pop}$. For ages 25, 45 and 90 the observed number of deaths is lower than expected in most years, which corresponds with the results, i.e. the posterior mean of the experience factor being smaller than 1. Analogously for age 65, the observed number of deaths is higher than expected in most years, which corresponds with the posterior mean of the experience factor being greater than 1.

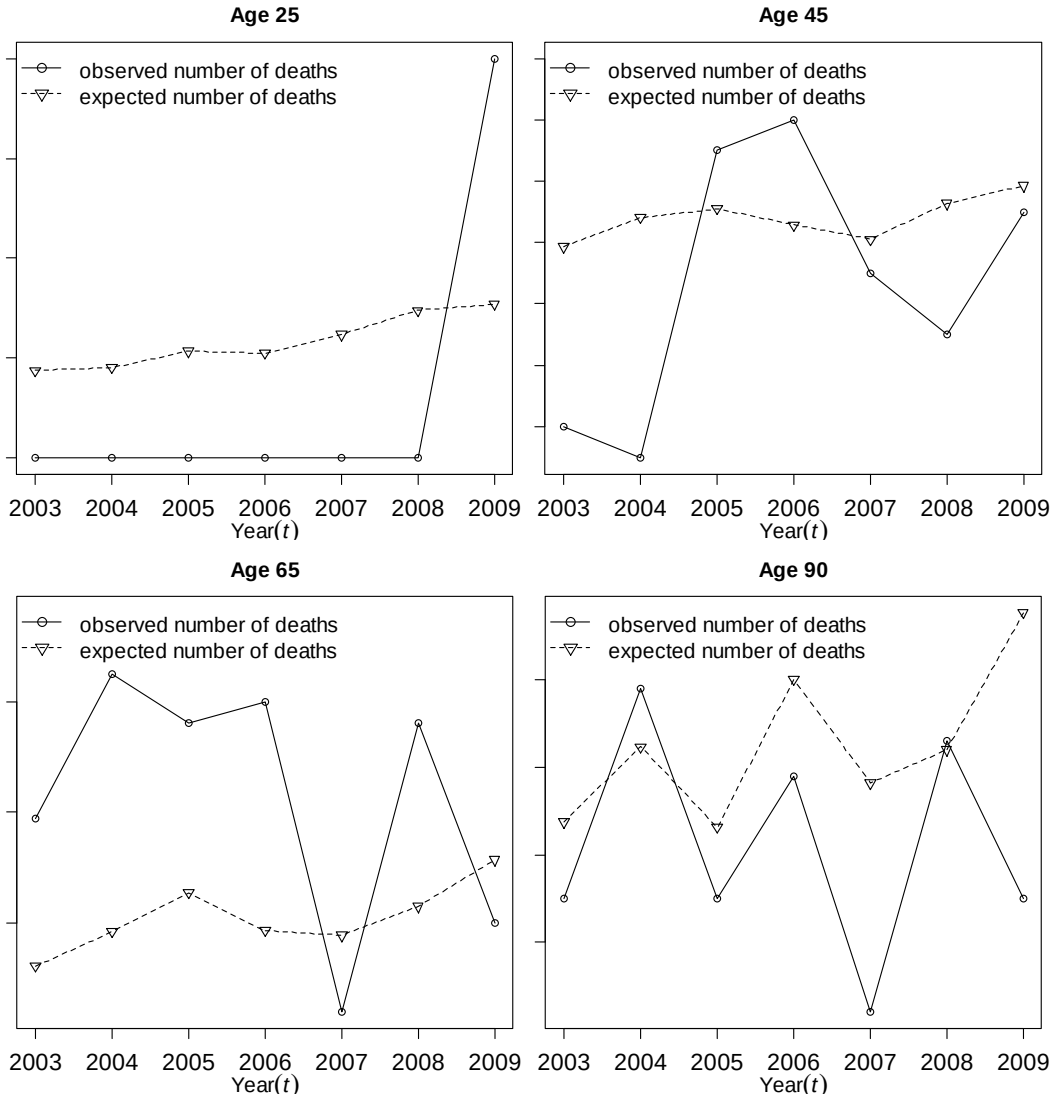

Figure 7.1: The observed and expected number of deaths for females

7.3 Sensitivity to the prior parameter

The parameters of the posterior distribution are sensitive to the parameter $\bar{\varphi}_x$ of the prior distribution as described in section 6.4. Therefore we perform a sensitivity analysis where the mean of the posterior distribution of the experience factor is calculated for different values of $\bar{\varphi}_x$. Figure 7.2 depicts the results of the sensitivity analysis and also shows the current values of $\bar{\varphi}_x$. For the ages 45, 65 and 90 we conclude that the posterior distribution of the experience factor is highly sensitive to the prior parameter $\bar{\varphi}_x$ for values lower than 0.05. For higher values the posterior distribution quickly becomes insensitive. For the age 25 we conclude that the posterior distribution remains sensitive to the prior parameter $\bar{\varphi}_x$ for a large range of values. Only for values higher than 1 the posterior distribution starts to become insensitive. From the estimated low values of $\bar{\varphi}_x$

we conclude that the Bayesian estimates of the portfolio experience factor are sensitive to changes in $\bar{\varphi}_x$ and that the portfolio data are not strongly informative with respect to $\bar{\varphi}_x$.

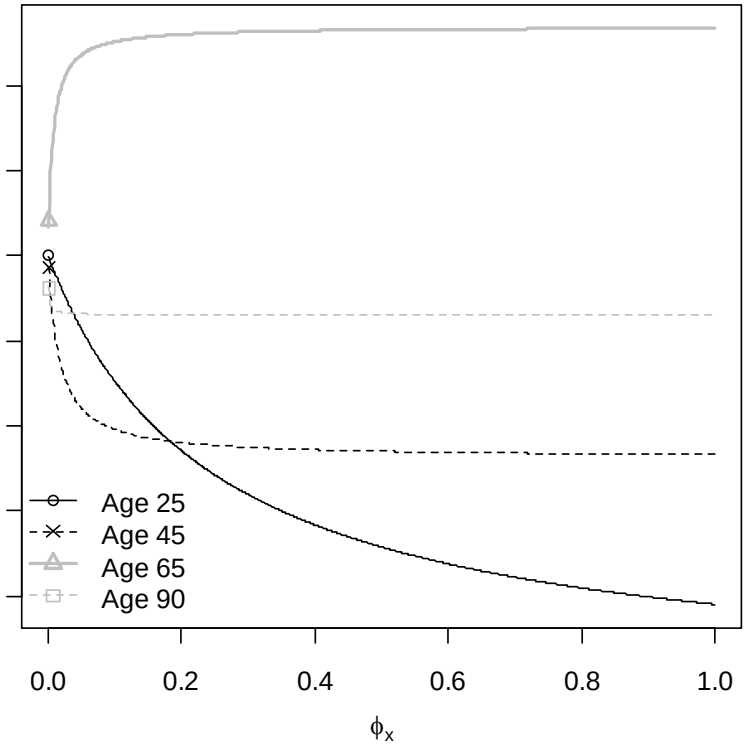

Figure 7.2: Sensitivity analysis to $\bar{\varphi}_x$ for females.

7.4 Results of the portfolio mortality forecast

In chapter 5 we obtained the Poisson-gamma Lee-Carter model fitted to the Dutch country population mortality. In section 7.2 we obtained the Bayesian estimates of the experience factor for the Dutch life and pension portfolio of AEGON. Using these two results we forecast the future portfolio mortality rate for the next 50 years, i.e. for the years 2010 to 2059, using the forecasting ability of the Lee-Carter model according to the approach described in section 6.6.

Figure 7.3 shows the results of the portfolio mortality forecast for females of age 25, 45, 65 and 90. For comparison the figures also contain the future country population mortality forecasted with the Poisson-gamma Lee-Carter model. Table 7.2 contains the average relative increase in width of the interval forecasts of the portfolio compared to the country population over the projection years, which is calculated same as has been done in section 5.5. We conclude that the more the Bayesian estimated portfolio mortality rate differs from the country population mortality rate, the wider the interval forecasts become.

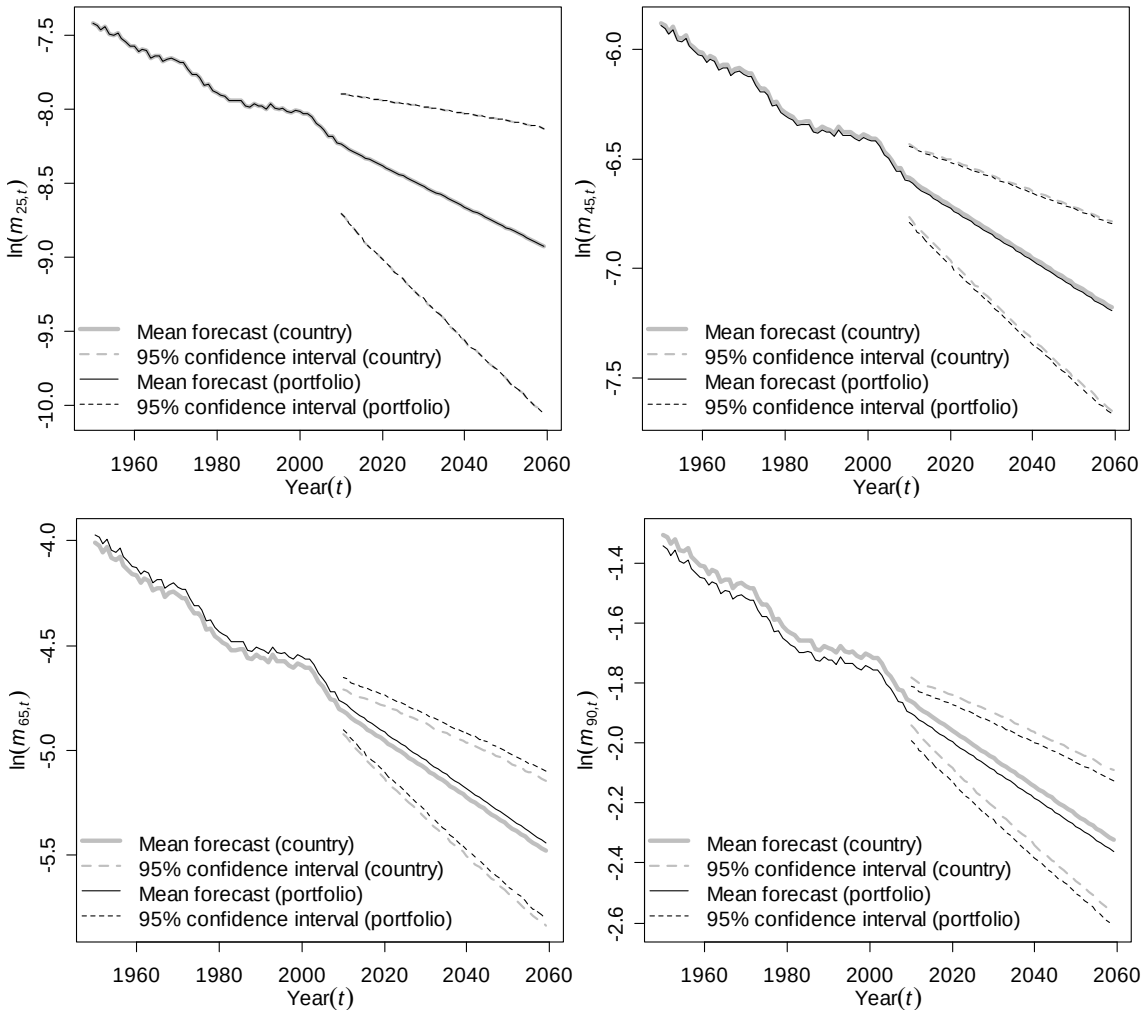

Figure 7.3: Portfolio and country population mortality forecast for females

	Average relative increase in width of the interval forecasts
Age 25	0.00%
Age 45	2.29%
Age 65	3.86%
Age 90	4.17%

Table 7.2: Average relative increase in width of the interval forecasts of the portfolio compared to the country population for females.

Chapter 8

Conclusions

In the first part of the thesis we investigated the applicability of the Lee-Carter model in a Poisson-gamma setting compared to the Poisson approach. More specifically, we looked into the Poisson-gamma approach with general dispersion parameter and with age-specific dispersion parameters.

The parameter estimates obtained by fitting the three models to Dutch population mortality data were in line with each other. Additionally, the dispersion parameters estimated by the Poisson-gamma models indicated the presence of overdispersion, i.e. heterogeneity, in the data. Compared to the Poisson approach both Poisson-gamma models delivered a significantly better fit, and the approach with age-specific dispersion parameters gave the significantly best fit.

We compared the forecast quality of the Poisson model to the Poisson-gamma model with age-specific dispersion parameters using simulation. The latter approach explicitly allows for overdispersion by introducing dispersion parameters and therefore can capture more variability. Due to this property the Poisson-gamma model yields wider confidence intervals than the Poisson model does.

Both the Poisson and Poisson-gamma model gave disappointing results in the backtest. The cause of this weak performance lies in the fast mortality improvements experienced in the recent years, in particular by the male population. The discrepancies between the mortality forecast published by the Dutch Actuarial Association, i.e. the "AG Prognosetafel 2010–2060", and the Poisson and Poisson-gamma model can be attributed to the same cause as well. On the contrary to the models under our investigation, the model underlying the "AG Prognosetafel 2010–2060" does distinguish a short term and a long term trend for forecasting mortality.

In the second part of the thesis we contributed to the research of stochastic models for portfolio mortality by proposing a Bayesian extension to the Poisson-gamma Lee-Carter model with age-specific dispersion parameters. Essentially, the *a priori* country population mortality rate gets adjusted to the *a posteriori* portfolio mortality rate using historical portfolio data and Bayesian inference techniques. The amount of heterogeneity, i.e. reliability, of the country population mortality data determines the sensitivity of the adjustment to historical portfolio observations. Additionally, the extension allows using the full forecasting ability of the Lee-Carter model for portfolio mortality. The

mathematics behind the extension is closely related to credibility theory more commonly applied in the field of non-life insurance.

We examined the Bayesian extension for portfolio mortality using the Dutch life and pension portfolio of AEGON. The estimated portfolio mortality rates did not differ significantly from the country population mortality rates. We conclude that the historical portfolio data have little impact on the *a posteriori* portfolio mortality rate. The cause of the insensitivity lies in the low value of the estimated dispersion parameters indicating a low amount of heterogeneity in the country population mortality data, and hence are considered reliable. Furthermore we conclude that the historical portfolio data are not strongly informative enough with respect to the country population mortality data.

Bibliography

Anon., 2010. *Generatietafels Pensioenen 2010*. Verbond van Verzekeraars. http://www.verzekeraars.nl/sitewide/general/nieuws.aspx?action=view&nieuwsid=870.

Anon., 2010. *Prognosetafel 2010-2060*. Utrecht: Actuarieel Genootschap. http://www.ag-ai.nl/view.php?Pagina_Id=333.

Bell, W.R., 1997. Comparing and assessing time series methods for forecasting age-specific fertility. *Journal of Official Statistics*, 13(3), p.279–303.

Booth, H., Maindonald, J. & Smith, L., 2002. Age-time interactions in mortality projection: applying Lee-Carter to Australia. *Working Papers in Demography*.

Bowers, N.L. et al., 1997. *Actuarial Mathematics*. 2nd ed. Society of Actuaries.

Brillinger, D.R., 1986. The natural variability of vital rates and associated statistics. *Biometrics*, 42, pp.693-734.

Brouhns, N., Denuit, M. & van Keilegom, I., 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal*, 2005(3), pp.212 - 224.

Brouhns, N., Denuit, M. & Vermunt, J.K., 2002a. A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, 31(3), pp.373-93.

Brouhns, N., Denuit, M. & Vermunt, J.K., 2002b. Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries*, 2, pp.105-30.

Cairns, A.J.G., 2000. A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics*, 27(3), p.313–330.

Carlin, B.P. & Louis, T.A., 2000. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall / CRC.

Chia, N.C. & Tsui, A.K.C., 2003. Life Annuities and Compulsory Savings and Income Adequacy of the Elderly in Singapore. *Journal of Pension Economics and Finance*, 2, p.41–65.

Delwarde, A., Denuit, M. & Partrat, C., 2007. Negative binomial version of the Lee–Carter model for mortality forecasting. *Applied Stochastic Models in Business and Industry*, 23(5), p.385–401.

Denuit, M., Maréchal, X., Pitrebois, S. & Walhin, J.-F., 2007. *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*. John Wiley & Sons.

Ghosh, J.K., Delampady, M. & Samanta, T., 2006. *An Introduction to Bayesian Analysis: Theory and Methods*. Springer.

Gilks, W.R., Richardson, S. & Spiegelhalter, D.J., 1996. *Markov Chain Monte Carlo in Practice*. Chapman & Hall / CRC.

Goodman, L.A., 1979. Simple Models for the Analysis of Association in Cross-Classifications having Ordered Categories. *Journal of the American Statistical Association*, 74(367), pp.537-52.

Heath, M.T., 2002. *Scientific Computing: an introductory survey*. 2nd ed. MGraw-Hill.

Hollmann, F.W., Mulder, T.J. & Kallan, J.E., 2000. Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100. *Working Paper 38, U.S. Bureau of Census*.

Kaas, R., Goovaerts, M., Dhaene, J. & Denuit, M., 2008. *Modern Actuarial Risk Theory*. 2nd ed. Springer.

Koissi, M.-C., Shapiro, A.F. & Hognas, G., 2006. Evaluating and extending the Lee–Carter model for mortality forecasting: bootstrap confidence intervals. *Insurance: Mathematics and Economics*, 38(1), pp.1-20.

Lawson, C.L. & Hanson, R.J., 1974. *Solving Least Squares Problems*. Prentice-Hall.

Lee, R.D. & Carter, L.R., 1992. Modeling and Forecasting U. S. Mortality. *Journal of the American Statistical Association*, 87(419), pp.659-71.

Lee, R.D. & Miller, T., 2001. Evaluating the performance of the Lee-Carter method for forecasting. *Demography*, 38(4), p.537–549.

Li, N. & Lee, R., 2005. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*, 42(3), p.575–594.

Li, S.H., Mary, R.H. & Tan, K.S., 2009. Uncertainty in Mortality Forecasting: An Extension to the Classical Lee-Carter Approach. *ASTIN Bulletin*, 39(1), pp.137-64.

Olivieri, A. & Pitacco, E., 2009. Stochastic Mortality: The Impact on Target Capital. *ASTIN Bulletin*, 39(2), pp.541-63.

Renshaw, A.E. & Haberman, S., 2003a. Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33, pp.255-72.

Renshaw, A.E. & Haberman, S., 2003b. Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projection. *Applied Statistics*, 52, pp.119-37.

Renshaw, A.E. & Haberman, S., 2006. A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38, p.556–570.

Renshaw, A.E. & Haberman, S., 2008. On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee–Carter modelling. *Insurance: Mathematics and Economics*, 42(2), pp.797-816.

Rubin, D.B., 1981. The Bayesian Bootstrap. *Annals of Statistics*, 9(1), pp.130-34.

van Duin, C. & Garssen, J., 2010. *Bevolkingsprognose 2010–2060: sterkere vergrijzing, langere levensduur*. The Hague: Centraal Bureau voor de Statistiek (Statistics Netherlands).

Wilmoth, J.R., 1993. *Computational Methods for fitting and extrapolating the Lee-Carter model of mortality change*. Technical Report. University of California Berkeley.